Made and Printed in Great Britain by
GREYCAINES
(Taylor Garnett Evans & Co. Ltd.)
Watford, Herts.

CONTENTS

ACKNOWLEDGMENTS

I WISH to record my thanks to the Editor of this Series, Professor E. A. Milne, for making several valuable suggestions for clarification when he read the manuscript. It should be stated here that Professor Milne has himself made major contributions, both of a pioneering and of a technical character, to much of the work described in this book. It is only because the book is primarily neither historical nor technical that it has to omit reference to such contributions.

I wish also to thank Miss P. Chapman and Miss J. G. Gardiner, of Royal Holloway College, for the great care with which they have read the proofs, and to Miss Gardiner also for checking my calculations.

W. H. McC.

June—1950.

CHAPTER I

INTRODUCTION

THERE are in the heavens certain objects, mostly inconspicuous but of vast importance, commonly known as "spiral nebulæ". For the moment we leave these out of account. Then we can assert that everything to be seen in the sky is found to pertain to one great system called the Galaxy.

It is estimated that the Galaxy contains several hundred thousand million stars. The Sun is one of these and may fairly be regarded as an average specimen. The stars form a lens-shaped distribution in space, or, as has been said in more homely language, their distribution resembles that of the currants in a bun. The Sun, attended by its planetary system, occupies a position near the "plane of the Galaxy" (i.e. where the butter is spread in the bun). That is why, since the majority of the stars are concentrated towards this plane, when viewed from our position close to the Sun they form the belt right round the sky which we know as the Milky Way. Though the Sun is near the plane of the Galaxy, it lies well away from the centre of the Galaxy, in fact probably nearer to the "rim" than to the centre.

The stars are not scattered at random through the Galaxy; they tend rather to be organized in a hierarchy of systems within the one great system. To begin with, about half of them possess companion stars and so form binary systems. A lesser proportion belong to systems of several stars which are near together compared with the average separation of all the stars in the Galaxy. Then there are "clusters" of stars, some irregular and diffuse like the Pleiades and containing from a few hundred to a few thousand stars, others more regular and compact called "globular clusters" containing tens of thousands of stars. Finally, viewing the Galaxy in the large, there are regions in which stars and systems of stars tend to be relatively concentrated and there are other regions which they tend to

avoid, so that the whole possesses a certain structural pattern, possibly ill-defined but nevertheless real.

The space between the stars is far more empty than the best vacuum that can be produced in a laboratory. But it is not utterly void. It is pervaded by an excessively tenuous distribution of *interstellar matter*, partly in the form of gas and partly "dust". This, too, is not uniformly distributed through the Galaxy. As a whole, it is concentrated towards the plane of the Galaxy and in addition it is broken up into clouds. Where the clouds are relatively most dense and best illuminated by starlight, they are seen directly as bright diffuse nebulæ like the well-known Orion Nebula. Where they are relatively dense but not illuminated, their presence is manifest by the obscuration of the background of stars as in the region known as the "coalsack" near the Southern Cross. But in general the existence of interstellar matter is recognized only by subtle influences upon the light coming through it from the stars, or by much less obvious obscuration effects. Despite its extreme tenuity, interstellar matter forms a very important constituent of the Galaxy and its total amount may be more than the total amount of matter forming the stars themselves.

This summarizes the way in which the parts of the Galaxy are related to each other in position.

The parts possess also a variety of relative motions. The components of a binary system revolve about each other; in general, in any system of stars, each star is in motion relative to the system. Further, each system as a whole has some motion relative to neighbouring parts of the Galaxy. The interstellar clouds are also in relative motion amongst themselves and amongst the stellar systems. And lastly, as is now known, the entire Galaxy is rotating in its own plane (the rotation being not like that of a rigid body but being faster for the inner than for the outer parts). This motion shared by all parts of the Galaxy provides perhaps the most convincing justification for considering it as forming a single whole.

It is now regarded as established that every one of the objects referred to as "spiral nebulæ" is itself a system generally similar in size, mass and composition to the entire Galaxy. For this reason they are now more usually called *external*

galaxies (and in any case many of them do not show the spiral structure to which these objects owed their earlier name). Only two or three of these galaxies can be seen with the naked eye, but larger telescopes reveal the existence of literally millions of them. Although they are all comparable as regards their most fundamental characteristics, they exhibit several different forms of general structure, and probably differ from one another in the extent to which their material is condensed into stars. They occur in groups and clusters, but the genuinely large-scale distribution of galaxies appears to be uniform in space, the average distance between neighbours being about a hundred times the average diameter of a single galaxy. The galaxies appear to be receding from each other and so produce the phenomenon known as the "expansion of the universe".

Such, in brief outline, is the known astronomical universe. It is necessary to sketch in this outline in order to show the setting of the subject matter of this book in the vast domain of investigation of which it forms a part.

. . . . . .

This book is concerned with the study of the Sun and stars as *individual* astronomical bodies. The object is to discover their physical constitution, and also as much as possible about their evolution in so far as this depends on processes within the stars themselves and not upon their relationship to their environment.

Until fairly recently this study comprised practically the whole of the science of *astrophysics*. Nowadays that science has grown to include the study of the physical constitution of the entire Galaxy and also that of other galaxies. By convention, the study of the relations between the galaxies is usually distinguished at present as the science of *cosmology*. But the distinction is not an essential one and is likely gradually to disappear.

Whatever the ultimate scope of astrophysics may be, its foundation must always be the physics of the Sun and stars. This is because a star is in fact the smallest body of matter

in the universe that leads an effectively independent existence over a long interval of time.

The stars which we shall consider belong to the Galaxy. Moreover, the data which we shall use are given by observations of stars in parts of the Galaxy relatively near to the Sun, merely because more remote stars cannot be observed in sufficient detail. But there are very good reasons for believing that, so far as their physical constitution is concerned, they provide an adequate sample of all the stars in the Galaxy. For that matter, they are probably typical also of stars in all the galaxies.

. . . . . .

The aim of the book is not so much to catalogue the conclusions of modern astrophysics as to attempt to demonstrate the way in which these conclusions can be reached. Some of the conclusions themselves are in fact scientific commonplaces. But, while doubtless many people are aware, for instance, that the internal temperature of the Sun is estimated to be of the order of ten million degrees, they probably have the impression that the result depends unavoidably upon highly complicated mathematical theory. It is true that such a result happened to be got in the first instance with the aid of somewhat difficult mathematics. It is, however, a fairly simple matter to construct a sequence of general physical arguments which enable us to rediscover this result without any mathematics except a little ordinary arithmetic. The writer tries to present such arguments in the case of a number of important astrophysical results.

It is therefore hoped that the following pages will enable the reader to understand some of the physical arguments used by astrophysicists in the more immediate interpretation of their observations and the further trains of reasoning which lead on to important general conclusions. It is also hoped that the reader will find satisfaction in the solution of the problems such as he may find in a good detective story in which the clues and arguments are fairly presented so as to lead to convincing conclusions.

With such objects in view, the reader may scarcely need to be warned that some of the discussion is not altogether easy going. Nothing deserving to be called mathematics is used, and each step in the arguments is simple enough in itself. But the reader must be invited to follow somewhat sustained reasoning in order to reach the ultimate conclusions. He may therefore in places find it profitable to devote a first reading to discovering the nature of these conclusions and a second one to the details of their derivation. The following explanation of the plan of the book may also be of assistance:

From the standpoint of our investigations, a star is primarily a generator of radiation, and all that we know about it is learned from observations made on this radiation. Therefore, all our investigations have to be based upon the physics of matter and radiation and the relations between them. This forms the subject of the next four chapters.

We then commence the study of the Sun by summarizing the observational data and proceeding to see how the foregoing physical theory enables us to infer the structure of the outermost layers of the Sun. In the following chapter we see first that, using only the known mass and size of the Sun and the general properties of matter previously described, it is possible to infer a great deal about physical conditions throughout the Sun's interior. This in turn enables us to formulate in adequately precise terms the central problem of our whole subject—that of the Sun's mechanism of energy-generation. We discover that only nuclear transmutations involving the consumption of hydrogen can offer a solution and that apparently only one particular such process can satisfy all the requirements. This finally makes possible a re-statement in complete form of the problem of solar constitution, and the main results of its solution are quoted.

We proceed then to the rest of the stars and review the data concerning their leading characteristics and also describe the classification of the stars. Thereafter we recall the well-known empirical relations which subsist between the luminosity and mass, and between the luminosity and spectral class. We find it possible to exploit a "method of similarity" so as to predict theoretically the manner in which the characteristics

of a star depend upon its mass and chemical composition. By these means we reach an explanation of the chief features of the empirical relations and this affords considerable insight into the constitution of a majority of the stars. The most numerous stars which do not conform to these results are the excessively dense "white dwarfs", but we find that accepted atomic theory can account for their distinctive properties. This part of the subject closes with a statement of problems concerning "giant" stars, which still lack a complete solution.

The final chapter deals with variable stars, first with periodic variables and then with "new" stars. The latter are seen probably to reveal the fate of stars which have used up their hydrogen supplies and also, when they take the form of "supernovæ", to account for the synthesis of heavy atomic nuclei in the universe. They appear to bring us to the limit of what can be learned about the evolution of the stars within the scope of their treatment as isolated bodies—and this is the scope of the present book.

It remains only to add that, although we have much to say about the results of astrophysical observations, we shall not attempt to describe the instrumental means of observation except to refer, when necessary, to the principles which they utilize.

IT will not surprise the reader to find that a book on the *physics* of the Sun and stars has to have quite a lot of space devoted to physics as well as that devoted to the Sun and stars. This chapter and the next three are intended to provide a sketch of the pure physics upon which the astrophysical investigations, subsequently to be described, depend most immediately. These physical preliminaries are restricted to our immediate requirements.

Energy. We shall find that some of the most important problems of astrophysics are concerned with *energy*—the generation of energy inside a star, its transport through the star, and its escape from the star. We shall begin with a brief account of energy itself.

Energy can appear in many different forms, familiar instances being potential energy, kinetic energy, heat energy, radiant energy, electrical energy, chemical energy. These are all equivalent in the sense that an engine could be contrived which would either convert such energy into mechanical work or else convert mechanical work into such energy. Were the engine one hundred per cent efficient, the number of foot-pounds[1] of mechanical work, produced or expended as the case may be, could be taken as a measure of the amount of the energy. Thus energy in any form could be measured in foot-pounds. For practical purposes, various units are in use in various connections. There is, for instance, the British Thermal Unit, and the fundamental importance of energy is recognized by the Act of Parliament which requires our domestic gas and electricity bills to be calculated on our consumption of *energy* in these units. But for purely scientific work, energy

[1] When we lift a pound weight through a height of one foot, we do a definite amount of *work* which is called one foot-pound.

Let us consider, for instance, the production of electric light by a hydro-electric system. It starts with a waterfall; the potential energy of the water is converted into the kinetic energy of a turbine; a dynamo converts the kinetic energy into electrical energy; this is converted into heat energy in the filament of an electric lamp; this heat energy streams away in the form of radiant energy—the light which it is the purpose of the whole system to produce. The light is absorbed by the objects on which it falls and its energy is there dissipated as heat energy. Were we to measure the quantity of energy fed in at the start and that flowing away at the end (and at the various intermediate stages where in practice a certain amount will also be dissipated as heat) we should find the two quantities to be equal.

Incidentally, if we go further back and enquire where the energy of the waterfall came from, we soon find ourselves in the realm of astrophysics. For the heat energy that evaporated the water, that made the rain, that fed the waterfall, came from the Sun. And, so soon as we ask where the Sun's energy came from, we reach a central problem of our subject.

Now the Sun's energy reaches us as *radiation* and it is necessary to say a good deal about energy in this particular form before we say anything more about the Sun.

Radiation. All that we have hitherto known about the stars and almost all we have known about the Sun, we have learned from the light which reaches us from these bodies. In recent years it has been discovered that the Sun, and probably also the stars, emits detectable radio waves. These will certainly enable us to add to our knowledge. These waves are, however, in any case, radiation of the same physical nature as light.

[1] c.g.s. = centimetre-gramme-second.
1 foot pound = 13·6 million ergs (approximately).

that is an aerial. The particular type of aerial and what the listener does with the radiation which it receives depends upon his particular requirements. If he is the navigator of an aircraft he will want to find the bearing of the transmitter and for this he employs one particular type of receiving set. If, on the other hand, the listener wants to hear a concert he employs another type of set and he will be insistent about its giving true tones and avoiding distortion. The musical listener wants, in fact, to hear as fully and accurately as possible what is happening at the transmitting end.

Now a star radiates light waves. The first essential for the astronomer is a means of picking up some of this radiation, that is a telescope. Like the wireless listener, he may want to use the radiation merely to determine the direction of the transmitter, in this case the star. But the astrophysicist is more like the musical listener. While his telescope plays the part of the aerial, his spectrograph and its accessories play the part of the receiving set in analysing the radiation received. Like the musician, the astrophysicist has to be very careful about the quality of the analysis. He, too, aims at knowing as much as possible about what is happening at the transmitting end, that is, in the star.

The analogy goes further. In both cases it is not the radiation itself which is directly observed but some effect produced by it. In the wireless case, the radiation is made to produce sound; in the astronomical case, the radiation is (normally) made to produce a photograph. Moreover, in both cases there is the feature of amplification. The wireless receiver achieves this by the use of valves and actually feeding in additional energy. The astronomer usually employs what we may call amplification by accumulation: he allows the effect of the radiation on his photographic plate to accumulate during a sufficiently long exposure-time. But in some very modern methods of observation he replaces the photographic plate

by a photo-electric cell and then employs means of amplification actually fundamentally the same as in the wireless case.

The reason we have got what seems to be a rather successful analogy is that basically we have not a mere analogy, but two instances of the same thing. For the radiation from a star and that from a wireless transmitter are physically the same sort of radiation. They are both electro-magnetic radiation and differ only in wavelength. Anyone can remind himself by reference to the dial on his radio set that ordinary wireless waves have a wavelength of a few hundred metres, and that so-called short-wave transmission uses wavelengths of a few metres. If the wavelength is reduced to something between about a metre and a centimetre we get the sort of waves used for "radar". If it is reduced to anything between a few hundredths to about a ten-thousandth of a centimetre we get infra-red radiation, including heat radiation such as comes from a domestic radiator. Wavelengths in the range about 0·00,008 to 0·00,004 centimetre are those of visible light. Proceeding from this, we get in succession ultra-violet radiation, X-rays and gamma-rays.

A fundamental property of all electro-magnetic radiation is that it travels through empty space with the "speed of light", which is very close to 3×10^{10} centimetres a second,[1] whatever its wavelength. Consequently, if we divide this number 3×10^{10} by the wavelength in centimetres, we get the *frequency* of the radiation. The possibility of characterizing waves by their frequency is another thing which has been familiarized by the radio. One well-known station is listed as transmitting on "1,500 m. (200 kcs.)", i.e. the wavelength is $1{\cdot}5 \times 10^5$ cm., and the frequency is 2×10^5 cycles per second. The product of these two quantities is 3×10^{10} cm. per sec. in agreement with what has just been stated.

The whole frequency-range of the various kinds of electro-magnetic radiation mentioned above which have been produced or studied in the laboratory covers about sixty octaves. Only

[1] We follow the practice, doubtless familiar to most readers, of expressing numbers as multiples of powers of ten. Thus, 10^6 means the number given by 1 followed by six zeros, i.e. a million. We use also negative powers: thus 10^{-6} means one-millionth, or, written as a decimal, 0·00,000,1.

about one octave of this range is visible light. So we might very well exclaim upon our good fortune that the stars should radiate in those very frequencies that our eyes can detect. Had it been otherwise, how should we ever have noticed the existence of the astronomical universe?

It is indeed fortunate that we can use our eyes in astronomy, but we ought to make a different judgment as regards cause and effect. For, if we are to have eyes which we can use, they must respond to the only sort of radiation which is abundant in our surroundings. This is solar radiation or, more precisely, solar radiation in the frequency-range in which it reaches the Earth's surface with greatest intensity. Now the stars are bodies generally similar to the Sun so that a significant part of their radiation is expected to be in the same frequency-interval. Consequently, we are led to suppose that, *whatever* had been the frequency-interval for the predominant part of the Sun's radiation penetrating our atmosphere, the human eye would have evolved so as to be sensitive in *that* interval, i.e. so that this became necessarily the interval of visible light. We are further led to suppose that in these circumstances the stars would in fact be visible. This point of view is entirely consistent with actuality.

To translate this into terms of our wireless analogy, suppose an intelligent individual is shown a radio set for the first time and has its working explained to him. He will not remark upon the fortunate circumstance that transmitting stations should broadcast upon just those frequencies which can be received by the set! If it is a British-made set, he will conclude that, since it has presumably been developed chiefly for listening to the B.B.C., those are the principal frequencies which can be transmitted by the B.B.C. and can reach the set. He may further argue that roughly the same conclusions must apply to transmission from broadcasting stations in general, though he may expect on general grounds that the set will miss rather more of the programmes of foreign stations than of the B.B.C. In this use of the analogy, the set corresponds to the eye, the B.B.C. to the Sun, and foreign stations to the stars.

Resuming the general argument, we notice that this has to make mention of radiation penetrating the Earth's atmo-

sphere and so to recognize, what we know to be a fact, that the atmosphere may not be transparent to radiation of all frequencies. It is interesting to notice that it follows from the general line of reasoning that our atmosphere must be relatively transparent to whatever ranks as visible radiation. Of course, it does not follow that it is not transparent to *any* other radiation. But, if there is an important fraction of solar radiation which cannot penetrate our atmosphere, then it does follow that that fraction is likely to be outside the visible range of frequencies, i.e. in the ultra-violet or infra-red. For the human eye would have had no opportunity to become sensitive to that fraction of the solar radiation. Thus these very general considerations lead us to expect that the Earth's atmosphere may be specially opaque to frequencies at either end of the visible range. This is actually the case and difficult problems arise on account of it.

We therefore expect direct astrophysical observations to be largely confined to a frequency range not greatly different from that of visible light, because our reasoning has led us to believe that this is the range that satisfies the two conditions of belonging to that in which the stars emit a significant amount of their radiation and to that which can penetrate the Earth's atmosphere. This expectation is borne out by experience. But we may say at once that the interpretation of this experience indicates that we should learn much more of what we want to know if astronomers could make their observations (using, of course, photographic and other means) at observatories outside the atmosphere. Actually, a first step in this direction has been taken by American astronomers who have recently got observations from apparatus projected on rockets to heights of the order of 100 km. above the Earth's surface.

Finally, although only such a relatively small range of frequencies is concerned in direct astrophysical *observations*, we shall find that radiation in the whole range from gamma-rays to radio frequencies plays an essential part in various astrophysical *phenomena*.

Analysis of radiation. In the case of visible radiation, the eye associates with each wavelength or frequency a particular

colour, some shade of red for the lowest frequencies in the visible range and some shade of violet for the highest. Now if any ordinary light is passed through a prism it is split up into some range of colours. This splitting up may be regarded as the resolution of the original light into its component frequencies and as showing that any ordinary light is a mixture, or superposition, of radiation covering some range of frequencies, and is not radiation of just one single frequency.

The result of sorting out the radiation into its separate frequencies arranged in natural sequence is called the *spectrum* of the radiation. The instrument which is used to produce and to study the spectrum is called a *spectroscope* or *spectrograph* according to its method of employment. The essential component in this instrument is either a prism (or train of prisms) or else a diffraction grating.

We often think of a spectrograph as an instrument for measuring wavelengths. More fundamentally, however, it is an instrument that enables us to compare the intensities of the light in different parts of the spectrum. The light being admitted to a spectrograph is a stream of energy. The instrument spreads out the light over its spectrum. By suitable means we can measure the energy-flow, i.e. the intensity, in any desired part of the spectrum. We can then represent such measurements by a graph showing the intensity-distribution over the spectrum, or, in other words, a graph from which we can read off what percentage of the incoming energy belongs to radiation in any stated frequency-interval. This graph is called the intensity-curve or *energy-curve* for the radiation.

It is scarcely necessary to add that we may speak of the spectrum and energy curve of any radiation, whether it is in the visible range or not.

Temperature radiation. Throughout this book we shall have a lot to say about *temperature*. We shall always suppose a temperature to be measured from the absolute zero. The unit used for numerical values is a degree centigrade, and in this unit an absolute temperature is simply the ordinary centigrade temperature with 273 degrees added. Most of the temperatures which are of astrophysical interest are so high

that the numerical difference between the centigrade and the absolute value is not of much significance. But physical laws involving temperature take their most natural and simple form when stated in terms of the absolute temperature.

Consider now any hot body such as a hot-water bottle or a lump of red-hot iron. There is no mistaking the fact that it gives out radiation and that the radiation depends upon the hotness, i.e. the temperature, of the body.

In order completely to describe the emitted radiation we should have to specify the quantity of energy emitted per second per unit area of the surface of the body for every frequency-interval of the radiation. We shall suppose that suitable devices are employed which ensure that the body concerned is maintained at some uniform temperature T throughout its substance. Then a very important physical law is that the specification required is given by what is known as *Planck's formula*. We shall not quote the formula itself because its slightly complicated mathematical appearance might obscure the essential simplicity of its physical implications. But we do require to state some of these implications.

(1) The formula contains only the temperature T and universal physical constants. Therefore, in all its characteristics, *the radiation emitted per unit area by a hot body depends solely upon its temperature* and not upon its shape, size, composition, or any other property. Hence it is natural to call this type of radiation *temperature-radiation*. (It is also called *black radiation* or *black-body radiation* for a technical reason into which we need not enter except to say that it is *not* because the radiation looks black!)

Actually, some general restriction on the nature of the hot body is required in order to ensure the applicability of Planck's law. It will be enough to say here that the body must be sufficiently large and dense. We know, for instance, that the radiation from a neon lamp is peculiar to neon and thus provides an example of a hot substance *not* emitting temperature radiation. But the fact that it gives this special radiation depends upon the density of the neon gas being small: were we to compress

the gas to a sufficiently high density and maintain it at a uniform temperature T it would, in fact, emit temperature radiation corresponding to the value of T. Conformity to Planck's law is the *normal* behaviour of a hot body, and a body exhibiting this behaviour is called a *black body*. (The term is explained here merely because it is so well-established in the literature of the subject. A

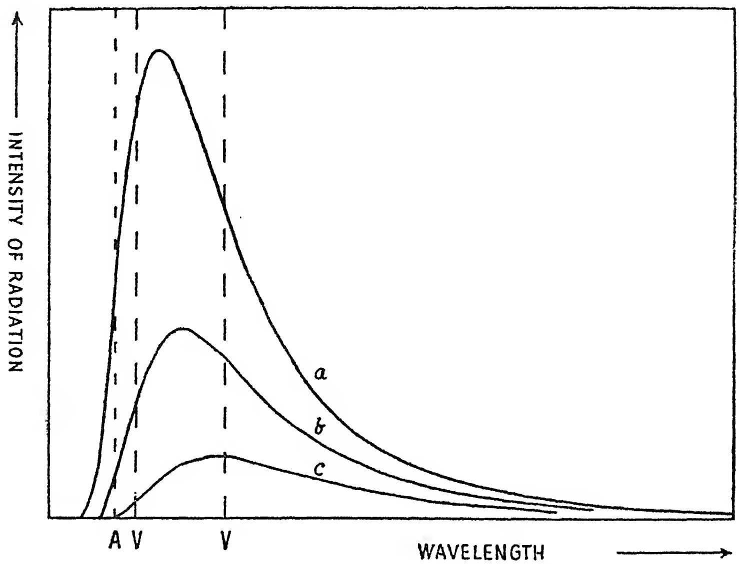


Fig. 1—Planck curves for (*a*) 6,000 degrees; (*b*) 5,000 degrees; (*c*) 4,000 degrees. V–V visible light. A = limit of atmospheric extinction.

body may be "white-hot" but may be described as a black body, this merely being a conventional way of saying that it is omitting temperature radiation *i.e.* that is behaving in accordance with Planck's law.)

(2) Planck's law is the mathematical expression for the *energy-curve* of temperature radiation. This we may call Planck's curve: it is of quite a simple appearance consisting of a single peak from which the intensity falls off (though not symmetrically) on both sides (see Fig. 1).

Moreover the peak intensity for any given temperature T occurs at a frequency which is proportional to T, this being the mathematical description of the property illustrated by the fact that a white-hot body is hotter than a red-hot body.

(3) Planck's law gives also, necessarily, the *total energy-flow* from unit surface-area of a hot body irrespective of frequency-distribution. This quantity may be called the total intensity of temperature radiation and is found to be proportional to the fourth power of the temperature, the constant of proportionality being a universal physical constant. The result is known as Stefan's law and can also be derived without a knowledge of Planck's law. It gives quantitative expression to the familiar fact of experience that the hotter a body the greater is its total emission of energy.

The immediate interest of these results for our purpose is that if a given body is known[1] to be at some uniform temperature T, then the value of T can be determined purely from observations of the emitted radiation. Roughly speaking, this means that we can regard a spectroscope as a "thermometer" which operates without material contact with the body whose temperature it measures.

Planck's law is such that this measurement of T can be made in various ways from observations of various characteristics of the radiation: a complete quantitative analysis of the radiation is not required. In the first place, we could use various measurements connected with what we should call the *quality* of the radiation. In practice, all such methods reduce essentially to observing some convenient stretch of the energy-curve and matching this with the corresponding portion of a Planck curve. There is a different Planck curve for every value of T and so the particular Planck curve which gives the match determines the value of T required. The

[1] This case is the simplest for our discussion. Under certain circumstances, the observations of the radiation can be used to *discover* whether the body is behaving as a black-body and then, if so, to determine its temperature.

importance of this sort of method is that it leads to the desired result without our having to know anything about the size of the emitting body or its distance from the observing apparatus.

In essence, the method is no more than this: Suppose we have one red-hot body whose temperature we *know* to be, say, 1,000 degrees. Then suppose we see another red-hot body glowing with a redness which precisely matches that of the first body. Then, whatever the size, shape or distance of the second body, we can assert that its temperature is 1,000 degrees. We have thus determined this temperature purely from the quality of the radiation. In the method described, for use in practice the match is made with a Planck curve corresponding to a known temperature instead of with the radiation from an actual body at a known temperature.

In the second place, we could alternatively use some measurement connected with the *quantity* of the radiation. We could measure the absolute intensity, in some frequency-interval, of the radiation received. Hence we could calculate, using geometrical considerations, the energy-flow from unit surface area of the body in this frequency-interval. The temperature for which Planck's formula predicts the same energy-flow is then the temperature required.

In contrast with the previous method, the geometrical considerations involved in the latter require that we should know the shape and size of the emitting body and its distance from the measuring apparatus. But now in any scientific work, and particularly in astronomy, if an observer encounters a method of measuring a quantity "X" that demands his knowing a quantity "Y", he automatically asks himself whether he can reverse the procedure and so obtain a method for measuring "Y" after finding "X" in some other way.

We have here an excellent example of this possibility. For, instead of using the second method to find T, suppose we already know it as a result of using the first method. Suppose we do know also the distance of the emitting body from the observing apparatus, and its shape, but that we do not know its size. Then, if we measure the intensity of the radiation as in the second method, we can calculate how big the body

must be in order to produce this intensity in accordance with Planck's law for the known temperature when it is situated at the known distance. Alternatively, if we know the size, we can infer the distance. In other words, we can regard a spectroscope and its accessories, not only as a "thermometer" for use when dealing with inaccessible bodies, but also as a "yard-stick" for measuring the size or distance of such bodies.

To put a spectroscope to such uses would generally be regarded as absurd in laboratory physics. But consider the problem of measuring the superficial temperature and radius of a star. Here we are entirely dependent upon the radiation emitted by the star. Also, as we have said, our own atmosphere permits the observation of this radiation only in a restricted frequency-interval. Finally, no direct measurement of the radius is in general possible, but astronomers have ways of measuring the distance. Consequently, we have just those conditions to which the present methods apply and in which they are essential. Actually, we have been describing in principle the simplest standard methods for determining stellar temperatures and radii.

We must, however, remind ourselves that these methods, as we have described them, apply only if we assume the star to radiate as a black body. It will be a matter for later discussion to examine this assumption. Owing to difficulties in regard to it, the present methods are not the most reliable even though they are the simplest.

CHAPTER III

STRUCTURE OF MATTER: MATTER AND RADIATION

In this chapter we first review some of the fundamental knowledge supplied by physics concerning the structure of matter. In the light of this knowledge, we then consider the processes by which matter and radiation interact. It is clear why we have to do this. For we want to know about the matter in the stars and all we can observe is the radiation they send us. So we hope to learn about the matter from an understanding of how it produces and influences the radiation.

Atoms. We have to consider the particles of which matter is composed. We start with the familiar fact that any specimen of any pure chemical element is composed of *atoms* of that element. The atom is the basic unit of the element as such. For many purposes such as the rudiments of chemistry and the kinetic theory of (monatomic) gases the atom may be regarded as a simple indivisible particle. But the discovery of the electron, more than fifty years ago, very soon led to the study of the *structure* of the atom and to our knowledge of the fashion in which the atom is composed of still more simple particles.

It is in these days common knowledge that an atom is a system consisting of a certain number of similar particles called *electrons* attendant upon another particle called the atomic *nucleus*. Each electron carries a definite negative charge of electricity and the nucleus carries a positive charge equal and opposite to the total charge of all the electrons in the complete atom. Further, the electrons are the same things as those whose movement along a wire constitutes an electric current in the wire. They are the things which enable a thermionic valve to operate a radio set. They play the key part in innumerable other physical phenomena and in all these

phenomena every electron is similar to every other.[1] In particular, the electrons in any one atom are exactly similar to those in any other, and they vary only in number from an atom of one chemical element to an atom of another. On the other hand, the nucleus of one element is essentially different from the nucleus of another. Also we know that the nucleus of any atom comprises nearly the whole of its mass.

The atomic nucleus is normally quite immutable, except for the natural decay of the few radioactive elements. But it is also common knowledge, and to most people extremely uncomfortable knowledge, that the nucleus *can* be disintegrated. For it is this fact which makes the atomic bomb such an alarming reality.

The spectre of atomic warfare, the social and economic changes that might result from the use of atomic power, and the beneficent clinical uses of certain artificially produced atomic nuclei today feature prominently in conferences, on the radio, in newspapers, magazines and books. All this has at least one repercussion for the writer of a book such as this. For it means that much of the terminology of atomic physics has now spread far beyond the preserves of professional science into the vocabulary of everyday use. As a consequence, the writer must suppose that the things themselves that he is writing about are not foreign to the reader's thought and language. He realizes, too, that much of what he says about these things is not unfamiliar even to the general reader. Nevertheless, it is desirable to set down here certain possibly well-known physical results since they are so immediately required in the subsequent astrophysical discussion.

It is evident from what has already been said that atomic physics can be divided into two parts. The first is that in which we can treat atomic nuclei as permanent particles. The second, which we may distinguish as "subatomic" physics, is

[1] We use the term *electron* (without qualification) for the particles to which these statements apply. Some readers will know that there are other sorts of particles sometimes called "positive electrons" and "heavy electrons". These are irrelevant to the topics of this chapter, and we shall have little occasion to mention them in this book. They are not "electrons" in the sense in which we use the term.

that in which we study the transformations of nuclei. Both are important in astrophysics. The first, however, is certainly all that is needed for the immediate interpretation of astrophysical observations, and it also goes a long way towards enabling us to draw the necessary inferences about conditions in astronomical bodies. The second is needed only when we come to study the sources of stellar energy which are operative under these conditions and the associated problems of stellar evolution. In the rest of this chapter we keep well within the first part.

Electron. We have spoken of an electron as a "particle" and so it is in the sense of being something got by the subdivision of matter. But the term "particle" has other connotations which imply that anything which it describes should at any instant possess a definite position, velocity, mass and, perhaps, shape, size and so on. Many readers will know, however, that an electron sometimes exhibits properties more akin to those of a "wave" than a "particle". Electrons passing through or being reflected by a crystal lattice, for instance, produce interference and diffraction effects closely analagous to well-known effects produced by light-waves. When this was first discovered in 1925, physicists started speaking of the "dual" character of the electron. This practice is still useful in any descriptive account of its behaviour.

The situation can be described with fair accuracy as follows:

If the observation of an effect immediately produced by an electron be called "seeing" an electron, then whenever we see an electron we see a "particle". That is to say, the observed effect—a scintillation on a screen, a track in a cloud-chamber or on a photographic plate, or whatever it may be—is of a sort that we should instinctively ascribe to the agency of a particle. But if we want to predict where an electron will be seen under given circumstances, then in the reasoning required for the prediction, we must treat the electron as a "wave". (Of course, in simple cases, "particle" reasoning may give the same answer as "wave" reasoning).

Even more simply, if somewhat less accurately, we may describe the experimentally established behaviour of the electron by saying that the electron behaves so as to exhibit

at the same time the properties of a particle and of a wave (or vibration) as closely as it is possible for these properties to be consistent with each other.

Quantum states. A consequence of this dual nature is that an electron in any isolated system, such as an atom, must be in one or another of a certain set of discrete states known as *quantum states*. We can see this from the following argument:—

Were the electron to behave purely as a particle, so that we could treat it in accordance with the ordinary ideas of *mechanics*, it could have any one of a continuous gradation of motions. But in any one of these states of motion it would, of course, possess a perfectly definite energy.

Were the electron, on the other hand, to behave purely as a wave or vibration, so that we could treat it according to the ordinary ideas of *wave theory*, then it could vibrate only in certain particular ways. For, by hypothesis, it is in an isolated system, and we know that a wave or vibration in an isolated system can possess only certain discrete frequencies. The free vibrations of a violin string, which is a simple instance of such a system, correspond only to one fundamental note and its harmonics. But in such a system any one of the permitted frequencies can be excited with arbitrary intensity, i.e. energy.

An actual electron, however, does its best to behave in both ways at once. So the state of motion in the "particle" description and the state of vibration in the "wave" description must be different aspects of the same thing which we may call simply the *state* of the electron.

But it at once becomes clear that the two descriptions can be applied to the same thing only if both are subject to severe restrictions. Since only certain states of vibration are permitted, we must conclude that *only certain states of motion are permitted*. Again, since a state of motion must have a definite energy, we must conclude that a permitted state of vibration can be excited *only with a definite energy*. Combining these conclusions, the electron can have only certain *discrete permitted states and permitted energies*.

We can express this last result by saying that the energy of the electron must be *quantized*. The same general conclusion

must apply also to a system of any given number of electrons. Thus it must apply to the whole electronic system of an atom. Therefore an atom can store energy only in certain stock amounts characteristic of its. quantum states. It is convenient to call these the energy-levels of the atom. The lowest level is called the "ground-state" and any higher level an "excited state".

Finally, we notice that the arguments we have been using would not apply to an electron that does not belong to any isolated system—what we should call an ideally *free* electron. Just as a violin string that is not attached to a violin can have no characteristic states of vibration, so the states of a free electron are not quantized. We may add that an electron which is not free is said to be *bound.*

Quantum Theory. Arguments such as those of the last section give some valuable physical insight and enable us to infer a number of general properties. But they are not rigorous and also they do not produce quantitative results. The physical theory which gives a proper quantitative account of the phenomena under discussion is the quantum theory. In the same way as ordinary mechanics has been developed to deal with ordinary particles and ordinary gross matter, and ordinary wave theory has been developed to deal with ordinary vibrations and such subjects as the propagation of light-waves, so quantum theory has been developed to deal with electrons and other fundamental particles and related problems concerning radiation. This theory deals from the outset with electrons as they are inferred from experiment to behave. It is a theory *sui generis* and not a combination of certain parts of mechanics and wave theory. But, because of the possibility of describing its results in terms of the dual nature of the electron, it is appropriately called *wave mechanics* (or *quantum mechanics*).

Quantum theory gives then, at any rate in principle, the means of calculating the properties of any electronic system, in particular the energy levels of any atom. In any particular instance, the required calculations may be too complicated to carry through completely, but in such cases the results can usually be got by semi-empirical means.

There is one general result which must be quoted. Any

quantum state corresponds to a characteristic frequency ν of the electron system regarded as a wave system. It possesses also a definite energy E. The theory shows that the two are related by the simple formula

$$E = h\nu, \qquad (1)$$

where h is a universal physical constant known as Planck's constant. (The energy E must be measured from some zero-point, and, of course, the formula is true only after this has been suitably chosen.)

Atoms. It is the physicist who studies the structure of the atom but it is the chemist who bestows the name of the element to which it belongs. Now a fact of the utmost importance, which has been established by the joint efforts of physicists and chemists, is that the chemical reactions of an element depend entirely upon the electronic system of its atom. These reactions are all that the chemist has to tell him that it is one element and not another. Therefore, if two complete atoms each possess the same number of electrons, the chemist must necessarily regard them as atoms of the same chemical element.

Let Z be the number of electrons in a complete atom of some element, and let us call the electric charge of an electron a unit (negative) charge. Then, as we have said, the nucleus of this atom must carry an equal and opposite charge, i.e. Z units of positive charge. In fact, the atom possesses Z electrons *because* it requires just that number to neutralize its nuclear charge. This makes it clear that the chemical classification of an atom depends upon its nucleus or, more precisely, upon the number Z belonging to the nucleus.

Thus an element is as completely identified by its Z-value as it is by its name. This is called the "charge-number" or "atomic-number" of the element.[1] All the elements are known

[1] It follows that it is only the electrical forces in an atom that are significant for the properties which we are now discussing. We need not say anything at this stage about the mass of the nucleus except to remark that these properties do depend ultimately upon the bare fact that the nucleus is much more massive than the electrons. But, so far as these properties are concerned, the exact value of the mass affects only the so called "hyperfine" features which do not concern us. We shall have to consider other properties for which the mass is of prime importance.

having atomic-numbers between $Z = 1$ (hydrogen) and $Z = 98$ (californium).

To take a particular case, the atomic-number of iron is 26. A complete atom of iron possesses, therefore, 26 electrons. The question we have now to consider is, What are these electrons doing? The answer which was current twenty or more years ago was that the atom is like a miniature solar system with the nucleus as the "sun" and the electrons revolving round it as "planets", the electrical attraction between the nucleus and the electrons playing the part which gravitational attraction plays in the actual solar system.

The interpretation of atomic phenomena provided by the quantum theory has shown this analogy to be remote. This is owing to the "dual" nature of the electron mentioned above. But, to the extent to which the electron behaves like a particle, the analogy is roughly valid and still affords a useful way of thinking about an atom. Also it suggests some terms in which it is convenient to describe the behaviour of the system, provided we do not read too much into them. Thus we shall sometimes speak of the "orbit" of an electron in the atom. This is to be taken to imply that what the electron is doing in the atom has some general similarity to a particle describing an orbit in a field of force, in particular that the electron has a definite total energy associated with its state.

In so far as we can use the orbital picture we must, of course, recognize that in view of the quantization only certain orbits are permitted. The set of permitted quantum states of the whole atom, and the total energy of each state, can be regarded as known from quantum theory. (This is not affected in principle by the fact that in particular cases the numerical results may have been obtained by semi-empirical means).

Emission of radiation. Consider an isolated atom. The fact of its being isolated implies, as we have seen, that it is in a definite quantum state. Suppose this to be an excited state "A" and let its energy be E_A. There are only a finite number of states of lower energy. Let state "B" be one of these with energy E_B. The question then arises: Can the atom pass from state "A" to state "B"? If so, it must rid itself of an amount

surrounding the atom.

One of the most important results of quantum theory is that, if the transition from state "A" to state "B" does occur, then the radiation thereby emitted is *radiation of the single frequency* ν given by the simple formula

$$E = h\nu, \tag{2}$$

where h is again Planck's constant. As a result of formula (1) we may state this alternatively as

$$\nu = \nu_A - \nu_B, \tag{3}$$

where ν_A, ν_B are the characteristic frequencies of the states "A" and "B" of the electronic system of the atom.

Thus the radiation which an atom of a given element can emit is radiation of only a discrete set of frequencies. Each permitted frequency is, according to the rule expressed by (3), the frequency-difference of the frequencies of a pair of quantum states.

At first sight, one might be inclined to expect that an atom in a quantum state in which the electronic system vibrates with characteristic frequency ν_A must emit radiation of frequency ν_A. But an atom *in* a quantum state is in a state of fixed energy and so is not emitting anything. An atom can emit, if it emits at all, only by leaving its initial quantum state. Also it can cease to emit only by attaining another quantum state. Therefore it is with the transition between the initial and final states that the emission is associated, and it must depend upon both these states. The total emission can only be the energy difference between these two states. Its spectrum is determined by the further appeal to quantum theory which we have now made (or to experiment) and is found to be of the simplest possible sort, that is to say, the whole emission is in a single frequency. Moreover, that frequency is related in the simplest possible way to the two

hen it will emit radiation whose spectrum consists of bright "lines" at the permitted frequencies.[2] whole array of spectral lines is called the *line-spectrum* lement. This spectrum is determined by the electronic of the atom, which in turn is determined by the charge- of the nucleus. It follows that each element must its own characteristic line-spectrum.

is well-known, as is also the fact that a highly effective identifying a given element is to generate its line n. This may be done, for instance, by introducing a the given material into the flame of a bunsen-burner an electrical discharge and examining the flame or e with a small spectroscope. The element can usually gnized from two or three of its spectral lines. His-

e reader might then ask why this is apparently so different behaviour of simple vibrating systems with which he is d. For instance, a violin string vibrating with any permitted y emits sound of the same frequency. But a violin string sound is not a truly isolated system. It has to be in contact air in order to produce sound. If it is truly isolated in a hen it emits nothing! Moreover, if we investigate means of he string to change over to a different permitted frequency, lo in fact reach results closely analagous to the stated quantum al ones.

st readers will be familiar with the appearance of the . of ordinary white light as seen in a spectroscope: it is a all the colours of the rainbow running from violet at one e strip to red at the other. If the white light is replaced by . single frequency, giving say yellow light, then the spectrum nsist entirely of a single bright yellow line. The appearance the same as if the original coloured strip were covered over aque screen having just one straight crack or slit, transverse rip, at the position of precisely the right shade of yellow in Similarly, if the light admitted to the spectroscope consists ure of light of several isolated frequencies, then the spectrum as the corresponding isolated bright lines. Thus the term *line* used in the text arises simply from the appearance l in the spectroscope.

torically, it was the existence of the line spectrum which led to the discovery of the quantum theory of atomic phenomena.

Transition-probabilities. Let us suppose once again that the atom is initially in the excited state "A" and that there are several states of lower energy so that state "B" may be any one of these. Then the preceding section states what would happen if the atom makes the transition to any state "B". Nothing has yet been stated, however, concerning which transition *will* occur.

What the quantum theory does in fact say about this is on first acquaintance quite surprising in a fundamental physical theory. It denies the possibility of predicting what *will* happen. Instead, it asserts that there is a definite *chance* that a transition will occur spontaneously within a given time, and it evaluates this chance for each of the transitions from the given initial state to states of lower energy. This chance is known as the *transition-probability* of the relevant transition.

In the case of certain transitions the transition-probability is found to be zero (at any rate to a first approximation). Such transitions are said to be *forbidden.* Consequently, the ordinary line-spectrum of an element does not contain a spectral line corresponding to every pair of energy-levels, but contains only lines corresponding to pairs having non-zero transition-probabilities. The rules prescribing the lines which can occur in this way are called *selection-rules* and these are fairly elementary deductions from the theory.

The theoretical calculation of the actual values of the probabilities is in general more difficult, but they can usually be evaluated empirically. For the purposes of our discussion, they may be regarded as known quantities.

This occurrence of probability right at the foundations of physical theory is of very profound significance. But, so soon as we recognize the existence of quantum states, we can see that it is unavoidable. If we were allowed to regard all atomic changes as being brought about by continuous processes, then we should expect to be able to predict them exactly, and there would be no place for probability considerations. Actually, however, the changes take place by discontinous "quantum jumps", from one quantum state to another. If

observe it we violate its isolation.

Moreover, we can observe the atom only by causing it to interact with some other agency (*i.e.*, in general, radiation) and it can do this only by performing a quantum jump. Thus the attempted observation would destroy what it is intended to observe! Therefore the concept of quantum states is inconsistent with the concept of exact prediction. We must consequently have recourse to the concept of probability, and we are in fact presented with just the sort of situation with which that concept is designed to deal.

The philosophical aspects, however, will not concern us further. We have always to deal with atoms in very large numbers. So the probability that any single atom in state "A" will jump to state "B" can be treated as the same thing as the *proportion* of all the atoms in state "A" which will make the jump. Our position is like that of any efficient caterer in a public restaurant. He cannot predict whether any particular customer, faced with a choice between, say, beef and mutton, will order the one or the other. But his knowledge of his business enables him to say what proportion of all his customers will prefer beef to mutton.

The practical significance of transition-probabilities is easily appreciated. Were we, for instance, to measure the absolute intensity of a particular spectral line, then the knowledge of the transition-probability for that line would enable us to calculate, for the gas emitting the radiation, the number of atoms in the relevant excited state (at any rate if conditions are such that re-absorption of the radiation is unimportant). Or again, were we to measure the relative intensities of spectral lines arising from several excited states, then the knowledge of the appropriate transition probabilities would enable us to determine the relative numbers of atoms in the various states. It is this kind of significance which is of astrophysical interest.

Exact reversal. Exceedingly general physical arguments show that if any particular atomic process is known to occur, then also the exact reverse of this process is something which can occur. This we shall call the principle of exact reversal.

If an atom in state "A" can jump down to state "B" with the emission of energy of amount E in the form of radiation of frequency ν (where $E = h\nu$), then this principle shows that an atom in state "B" placed in a field of radiation of frequency ν can jump up to state "A" with the *absorption* of energy, from the radiation, of amount E.

It is known also, from an extension of the principle, that the chance of the occurrence of the latter process can be calculated in terms of the intensity of the radiation and the transition probability for the original process.

Further, if the atom in state "B" is placed in a field of radiation of any frequency, which is *not* the frequency of the radiation emitted in some transition having "B" as the terminal state, then the atom can absorb no energy from the radiation. For there is no process of which such absorption could be the reverse.[1]

Radiation which contains all frequencies in some continuous frequency-range we shall call *continuous radiation*. The spectrum of such radiation is a continuous band of colour (comprising all the colours of the rainbow if the range includes the whole visible range) and its energy-curve is a smooth curve. Temperature radiation is an example of continuous radiation.

Suppose now that we have a "gas" of atoms of a given element including atoms in all possible quantum states, and suppose we interpose it in a stream of continuous radiation. The results just enunciated show what the effect must be. The spectrum of the radiation after passing through the gas will be interrupted by a set of dark lines at precisely the frequencies of the spectral lines of the element, owing to the absorption of the radiation of those frequencies. Since the gas can absorb no other frequencies, the spectrum will be otherwise

[1]Some readers may remark upon our not mentioning the possibility of the "stimulated emission" of quanta of radiation. For our purposes here, however, this phenomenon is unimportant.

Thus, as we might expect on general grounds, the reversal of the atomic processes produces a reversal of the spectrum. In the case considered earlier, we had bright spectral lines with darkness between them; in the present case we have dark spectral lines in the same positions with light in between. The former we call an *emission spectrum* and the latter an *absorption spectrum.*

In the case of an absorption spectrum we shall use the term "strength" of a line to denote the amount of the continuous background which is missing on account of the absorption line. Analogously to the emission case, this strength provides a datum for finding the number of absorbing atoms producing the line.

It is scarcely necessary to remark that not every spectrum, emission or absorption, produced by a given element will exhibit every possible spectral line of the element. The presence of any particular line depends upon the state of excitation of the atoms in the specimen of the element producing the spectrum.

Finally, it is clear that if the "gas" producing the spectrum is a mixture of various elements, then each of these elements will yield its own characteristic spectrum as though the other elements were absent.

The mere existence of line-spectra is an old story. The manner in which it has been recounted in this chapter is, however, intended to be preparatory to some subsequent discussion of less rudimentary features which are important in modern astrophysical investigations.

CHAPTER IV

ATOMIC PROCESSES: FORMATION OF SPECTRAL LINES

In the preceding chapter we described the emission and absorption spectra of a "gas", but we contemplated the "gas" as merely an aggregate of atoms each behaving as though in isolation from all the rest. This may indeed be regarded as a first approximation to an actual gas if we have to deal only with a sufficiently small quantity in a sufficiently rarefied state. But in cases of interest of an actual gas the atoms do not behave as though isolated from each other. Each is affected by the others both by material encounters with them and also by the radiation which they emit. In the present chapter we shall therefore consider the atomic processes which can occur and their effects upon the spectrum produced by the gas. It is these effects which render the spectrum capable of telling us what we are interested in about the gas—its temperature, density, extent, and so on.

Photons. Before proceeding with our discussion it is desirable to introduce another concept in regard to radiation. We have hitherto been treating radiation in the usual manner as a *wave* phenomenon. But it follows from what was said in the preceding chapter that radiation of any particular frequency ν can interact with matter only in definite quanta each having precisely the amount of energy $h\nu$. Since we can observe radiation only by its interaction with matter, we must therefore think of it as being always parcelled up in such quanta. This is, however, equivalent to thinking of radiation rather as a *corpuscular* phenomenon.

The situation is in fact generally similar to the one we met in the case of electrons. Like electrons, radiation possesses a dual nature. In its detailed interaction with matter it behaves as though composed of particles, or *photons* as they are called.

of photons was introduced by Einstein and Bohr and led to Bohr's theory of atomic structure in 1913. The "wave" nature of the electron was a concept introduced by de Broglie in 1924 and was soon shown, chiefly by Schrödinger, also to result necessarily in the existence of quantum states.

While there is the similarity referred to between electrons and photons, there is also, of course, ample distinction between them. An electron carries a fixed electric charge and can travel through empty space along any path (depending upon its mode of projection and the electromagnetic field in the space). A photon carries no charge and can travel through empty space only in straight lines and only with the speed of light. An electron may become bound to a nucleus but it retains its individuality. If the resulting atom captures a photon, then the photon as such is annihilated, only its energy (and momentum) remaining as an increase in the energy (and momentum) of the atom.

Ionization. We may return for a moment to the "solar system" model of an atom. In the actual solar system, a planet remains in its closed orbit round the Sun, or, as we may say, remains "bound" to the Sun, because it does not possess enough energy to escape from the Sun's gravitational field. If by any means the planet could receive a sufficient access of energy, for instance by colliding with some other body, it would effect its escape, or, as we may say, become "free" from the Sun. In an analogous way, the Z electrons in a normal atom of atomic number Z are all "bound" to their system. But if sufficient energy is communicated to one of the electrons, then that electron will leave the atom and become "free". The atom lacking this one electron has then a net positive electric charge of one unit and is called a *singly ionized* atom. If it loses a second electron it is said to be *doubly ionized*, and

so on until it is Z-ply ionized and consists merely of the bare nucleus. Any ionized atom is called a positive *ion*.

Photo-ionization. Let I be the amount of energy required just to liberate the most loosely bound electron in some state of the atom. Then, so far as energy considerations alone are concerned, any photon of energy greater than or equal to I could effect the liberation. For we recall that the energy of an electron when free is not quantized. So *any* surplus energy of the photon over and above the amount I could be taken up as kinetic energy of the free electron.

This process of ionization by the agency of a photon is called *photo-ionization* and is in fact well known to occur in practice. The process may also be called a *bound-free transition* as compared with the quantum jumps previously considered which may be called "bound-bound" transitions.

The quantum theory supplies values of transition probabilities for bound-free transitions as well as for bound-bound transitions.

By the universal rule relating the frequency of radiation to the energy of its quanta, the frequency ν_0, say, of the radiation which is just able to ionize the atom in the given state is determined by $h\nu_0 = I$. The absorption of any quantum of frequency greater than ν_0 would ionize the atom. Hence the lines in the absorption spectrum of atoms in this state must all be at frequencies less than ν_0. But we have just seen that *any* quantum of frequency greater than ν_0 may be absorbed. Therefore the complete absorption spectrum consists not only of the sequence of absorption lines below frequency ν_0 but also of absorption at all frequencies above ν_0. This latter we call *an absorption continuum.* Actually, although it extends theoretically over all frequencies above ν_0, both theory and observation show that it fades out rather rapidly with increasing frequency.

The absorption spectrum of a gas containing atoms in various excited states must contain an absorption continuum corresponding to each of these states. Such continua are not very important in laboratory spectra and very few of them can be clearly identified in astronomical spectra. Nevertheless

physical theory has to do with the propagation of continuous radiation through a gas. Line absorption, which is what we are most familiar with in a gas, offers very little obstruction to the radiation. For it affects only the fraction of the radiation whose frequencies are exceedingly close to those of the absorption lines, and this is a very small fraction of the whole. So far as the rest of the radiation is concerned, the gas might as well not be there at all if the only obstruction it has to offer is due to line absorption. And indeed, the gases with which we most commonly have acquaintance, such as ordinary air, do appear to be almost perfectly transparent. But now, as soon as we recognize the possibility of photo-ionization we realize that we have a means by which the gas can obstruct every part of radiation. For, considering radiation of any frequency whatever, suitably excited atoms can be ionized by it and can therefore absorb it.

The reason ordinary air is so transparent to visible radiation is that it contains a negligible proportion of atoms (or molecules) which can be ionized by it. The air would be relatively opaque to radiation of very much higher frequencies, or alternatively, could be made opaque to visible radiation by making it very much hotter and so exciting its atoms into suitable states.

For the purpose of astrophysical theory we must in fact make ourselves aware of any agency which can cause appreciable obstruction to continuous radiation, and we shall see that there are two or three others besides photo-ionization.

Before leaving this particular phenomenon, however, we should remark upon its reverse. The reverse of the ionization of an atom owing to the absorption of a photon is the capture by an ion of a free electron with the emission of a photon. So the emission spectrum of a gas containing ions and free electrons will show continuous emission in the regions where the absorption spectrum shows continuous absorption.

Spectrum of ionized atoms. A singly ionized atom of atc number Z possesses $Z - 1$ bound electrons. Such atoms rise to a spectrum (emission or absorption) in all resp qualitatively similar to that of normal atoms of atomic nun $Z - 1$, which possess the same number of electrons. ' frequencies at which corresponding features occur are, howe different since the nuclear field is different. Similar res apply to further stages of ionization. Spectroscopists dea familiarly with spectra of ionized elements as with those normal elements and many such spectra are of interest astrophysicists.

Spectral frequencies. In all the foregoing discussion of ato spectra, nothing has been said about the actual magnitude the radiation frequencies occurring. Broadly speaking, ato excitation energies are such that most elements have so important spectral lines in the visible part of the spectru Many other important lines occur outside this part, partic larly in the ultraviolet. Quite a few atoms and ions commo represented in astronomical spectra give absorption contin in the visible region, though they may give more importa ones in the ultraviolet. The major part of the spectra multiply-ionized heavy elements retaining only a few their electrons is in the X-ray region, far beyond the ultr violet.

Free-free transitions. Since the energy of a free electron is n quantized, one might think that it could absorb any photo and convert the total energy of the photon into an increase : its own kinetic energy. If so, we should naturally call th process a *free-free transition.* It can, however, be proved th such a transition cannot occur unless the electron is acte upon by some external field. Free-free transitions are possib for an electron moving in the field of a positive ion (thoug not, of course, in this case bound to the ion) with the absorptio or emission of photons. As another possible agency of con tinuous absorption, such transitions have considerabl significance in astrophysics, though in general they are les important than bound-free transitions.

Scattering by free electrons. If a completely independent electron encounters a photon it cannot, as we have just said, absorb the photon. But this is not to say that there can be no interaction. In so far as the electron and photon may be regarded as particles, the result will be much like that of any collision between a pair of particles: their energies and directions of motion may be changed by the collision.

The change of energy need not concern us here. What does concern us is that, though the photon cannot be absorbed, it may be deflected. Hence, if we have a stream of radiation passing through a gas of free electrons, the electrons may obstruct its flow by "scattering" the radiation. That this effect does occur has been known for a long time and its magnitude can be calculated, and verified by experiment. It is often called "Thomson scattering", after J. J. Thomson, who first performed the calculation. In the very highly ionized material of the stars its effect may be paramount in controlling the flow of radiation.

Atomic collisions. So far we have only considered transitions produced by or producing radiation. The only other way they could be caused is by the direct action of matter upon the system performing the transitions. In a gas, the atoms, ions, and any other particles present are always in a state of thermal agitation as a result of which they are continually colliding with each other. The collisions are the only direct interaction with other matter experienced by the particles. Now it is in fact known that the collisions may produce transitions in one or both of the colliding particles. The effect may be to excite or ionize an atom (inelastic collision) or to de-excite an already excited atom (super-elastic collision). The energy of excitation or of ionization gained or lost by the atom, instead of being balanced by a change of radiant energy, is balanced by a change in the kinetic energies of the colliding particles.

This, incidentally, explains the most familiar method of producing excited atoms. For when we "colour" a bunsen flame with sodium, strontium, copper or whatever we please, the atoms of the colouring element are excited by collisions

Spectrum of ionized atoms. A singly ionized atom of atomic number Z possesses $Z - 1$ bound electrons. Such atoms give rise to a spectrum (emission or absorption) in all respects qualitatively similar to that of normal atoms of atomic number $Z - 1$, which possess the same number of electrons. The frequencies at which corresponding features occur are, however, different since the nuclear field is different. Similar results apply to further stages of ionization. Spectroscopists deal as familiarly with spectra of ionized elements as with those of normal elements and many such spectra are of interest to astrophysicists.

Spectral frequencies. In all the foregoing discussion of atomic spectra, nothing has been said about the actual magnitude of the radiation frequencies occurring. Broadly speaking, atomic excitation energies are such that most elements have some important spectral lines in the visible part of the spectrum. Many other important lines occur outside this part, particularly in the ultraviolet. Quite a few atoms and ions commonly represented in astronomical spectra give absorption continua in the visible region, though they may give more important ones in the ultraviolet. The major part of the spectra of multiply-ionized heavy elements retaining only a few of their electrons is in the X-ray region, far beyond the ultraviolet.

Free-free transitions. Since the energy of a free electron is not quantized, one might think that it could absorb any photon and convert the total energy of the photon into an increase in its own kinetic energy. If so, we should naturally call the process a *free-free transition.* It can, however, be proved that such a transition cannot occur unless the electron is acted upon by some external field. Free-free transitions are possible for an electron moving in the field of a positive ion (though not, of course, in this case bound to the ion) with the absorption or emission of photons. As another possible agency of continuous absorption, such transitions have considerable significance in astrophysics, though in general they are less important than bound-free transitions.

with the atoms or molecules of the burning gas. The function of the burning is merely to endow these particles with the energy required to render the collisions sufficiently violent. We may remark that the reason we obtain the characteristic colour, i.e. the emission spectrum of the element, is that under the conditions prevailing an excited atom has a good chance of emitting its energy as radiation before it suffers a further collision.

The effects produced by collisions are all calculable (at least in principle) using quantum mechanics and the dynamical theory of gases.

Molecules. This may be a suitable point at which to dispose of the question of molecular, as distinct from atomic, phenomena. Under terrestrial conditions, atoms are to be found more often than not in molecular combination with other atoms. In the universe as a whole, however, conditions are such that molecular combinations are exceedingly rare and play no part at all in the vast majority of the physical processes which take place. We may therefore disregard any possibility of molecular phenomena in all our general discussions.

State of a gas. The picture of a gas which we have to keep in mind is this:—The material particles are in chaotic motion and continually colliding with each other. Some collisions are sufficiently energetic to ionize atoms and consequently these particles comprise ions and free electrons as well as complete atoms. And here it should be stated that these particles must remain intermingled in such a way that any appreciable volume of the gas contains effectively equal numbers of positive and negative charges: electrostatic forces preclude any systematic sorting out of the particles of different types. Some collisions produce excited atoms or ions, and these may emit photons before suffering further collisions. Photons may also be produced in free-bound and free-free transitions. Therefore radiation *must* be present, whether or not any is entering from outside the gas. So the various radiational transitions which have been mentioned must be taking place.

Of the other extreme, in the case of, say, hydrogen in a stellar interior, the number of non-ionized atoms would be negligible and the radiation present might exert a pressure comparable with the gas pressure. The present discussion, however, is concerned with general principles and must therefore take account of all possible features.

Suppose now we have a gas of which we know the chemical composition, that is, we know what chemical elements are present and in what proportions. Suppose also we know everything we need to know about the *external* physical conditions to which the gas is subjected.

Then at any instant the gas must be in a certain state of excitation and ionization, the atoms, ions, and electrons must be in a certain state of agitation, and a certain amount of radiation must be present. Suppose we want to know this *internal* state of the gas.

More specifically, suppose we want to know:

for the atoms of each element, the proportions in the various stages of ionization,
for each of the types of atom and of ion, the fractions in the various quantum states,
for each of the types of particle, the average kinetic energy of the chaotic motion, and
for the radiation present, the total amount and its energy-curve.

In principle, all these features of the internal state of the gas are actually calculable from a knowledge of the quantum-states and transition-probabilities of the individual atoms and ions, for any given set of external conditions.

It may well appear from all that has been said that such a calculation would be prohibitively difficult. It would be like the following problem: For a country of stated size and total

population, knowing the habits and capabilities of individual human beings, to calculate the numbers of individuals engaged in various industries, being given only the figures for external trade.

To make this latter problem more precise, suppose there is no external trade but suppose it is given that the country is in a steady state of no statistical increase or decrease of its material reserves. It is moderately clear that the problem should in fact be soluble. For, unless the population were distributed over the various industries in a particular manner, each industry would be producing either too much or too little of its particular product for the consumption of the whole population. In principle, the requirements of the problem could be expressed mathematically and the resulting equations solved so as to give the desired information.[1]

Even if the mathematical solution of this problem were feasible, it would, however, be possible and much simpler to obtain the required information by merely taking a census of the population. In the problem of the gas, on the other hand, a census is not possible and the required information is to be had only as a result of the mathematical investigation. In the physically most important case the mathematical solution is fortunately well-known. We proceed to describe this case and then briefly to state how other cases can be treated.

Thermodynamic equilibrium. We consider a quantity of gas contained within an enclosure whose walls are rigid and impervious to matter and energy in any form. For the sake of simplicity and definiteness, we shall assume also that no forces such as gravitation act upon the gas. The gas must acquire a steady state in which its gross properties such as pressure and density are uniform throughout the enclosure and constant in time. The gas is then said to be in *thermodynamic equilibrium.*

The existence of the rigid enclosure is not essential for the

[1] The possible problem is here only sketchily indicated; its proper statement would demand much more detailed formulation. Somewhat analogous problems are in fact formulated and solved in the mathematical treatment of epidemiology.

to a gas (or any other physical system) if, and only if, it is in thermodynamical equilibrium. Furthermore, the statistical theory of gases shows that, for a given quantity of a given gas within the enclosure, the state is completely determined by the temperature T alone.

It is sufficient for the reader to know this and to know that the various features of the internal state which we have enumerated are all calculable from known formulæ in this case of thermodynamic equilibrium. But the following slightly more detailed statement of the results supplied by the statistical theory of gases may be of interest.

(a) The mean kinetic energy of a particle of the gas is $\frac{3}{2}kT$, where k is a certain universal constant not depending upon the mass, or other characteristics, of the particles.
(b) The state of excitation of any species of atom or ion depends, in accordance with a certain known formula, only upon the energy-levels and the temperature, and not upon the density. The formula gives the expected result that the proportion of excited atoms increases with increasing temperature.
(c) The degree of ionization of any atomic species depends, in accordance with a known formula, only upon its energy-levels, the temperature, and upon the density of that species and of the free electrons.
(d) The radiation present is temperature radiation, of such a sort that, were a small aperture cut in the walls of the enclosure, then the escaping radiation would be exactly the same as if that aperture were part of a body at temperature T emitting temperature radiation.

The point to notice about the results (b) and (c) is that the population of any permitted state is quite independent of the processes by which the particles concerned enter and leave

that state and of the relevant transition probabilities. This is not true of conditions other than that of thermodynamic equilibrium. It is this feature that renders the derivation of the required formulæ simpler for thermodynamic equilibrium than for other cases.

It can be shown that this feature is equivalent to the fact that in thermodynamic equilibrium any atomic process proceeds, statistically, at precisely the same rate as its exact reverse. Thus, for instance, the same number of atoms in a particular quantum state are de-excited per unit time by superelastic collisions as are excited into that state by inelastic collisions. This is true no matter what radiative transitions are affecting the same quantum state at the same time. This general property is called the principle of *detailed balancing*, and it is from it that our principle of exact reversal is derived.

Conditions other than thermodynamic equilibrium. In practice the condition of ideally perfect thermodynamic equilibrium is never realized. But conditions which, for practical purposes, are indistinguishable from it, are often encountered. The physical systems of interest in astronomy are, however, in general far removed from the thermodynamic equilibrium. Nevertheless the state of thermodynamic equilibrium provides the obvious *norm* with which to compare them and, when possible, in terms of which to describe them.

In this section we make a preliminary survey of the conditions to be considered later in their appropriate astrophysical contexts.

(a) Local thermodynamic equilibrium. In a star the temperature may vary from over ten million degrees near the centre to a few thousand degrees near the surface, whereas a system in thermodynamic equilibrium must, of course, have the same temperature throughout. It is nevertheless not difficult to infer that, though the star considered as a whole is not in thermodynamic equilibrium, yet the state of the material in any small region inside the star is exceedingly close to that of thermodynamic equilibrium at some particular temperature.

dynamic equilibrium (or very nearly so) at that temperature, because temperature is defined only for thermodynamic equilibrium. The system is then said to be in *local thermodynamic equilibrium*.

(b) Non-thermodynamic equilibrium. In a gaseous nebula in interstellar space the excitation of the atoms is produced almost entirely by the incident stellar radiation. The degree of excitation can be found only by calculating the numbers of atoms, entering and leaving the various quantum states in the given radiation field, using the appropriate transition-probabilities. Owing to the very low density, collisions between the atoms have no appreciable effect upon their excitation; in other words, the random motions of the atoms have no effect.

This is an example of a system which simply has no "temperature" in any strict sense. Nevertheless, the random motions of the atoms would be indistinguishable from the thermal motions at some particular temperature. This we could call the "kinetic temperature" and it would be what we should mean were we to speak, using somewhat loose terms, of "the temperature" of the interstellar material concerned. But, from what has just been said, the degree of excitation is virtually independent of the kinetic temperature in such a case.

(c) Pseudo-thermodynamic equilibrium. It can happen in certain cases that a gas is subject to external conditions which are manifestly not those necessary to produce thermodynamic equilibrium yet which conspire to produce a state that roughly imitates thermodynamic equilibrium at some particular temperature. A case in point is the material very close to the apparent surface of the Sun. This material is exposed on one side to the full glare of solar radiation and on the other side faces relatively empty space, conditions which depart

all resemble those in a gas at about 5,000 degrees. For want of a better term, we can describe the gas as being in *pseudo-thermodynamic equilibrium* at this temperature. In regard to detailed phenomena, the departure from true thermodynamic equilibrium may be all-important, but this description serves usefully to provide a rough general conception of the state.

(d) Quasi-thermodynamic equilibrium. It may happen that one particular atomic process in a gas proceeds excessively slowly by comparison with all the other processes. In such a case the system may attain thermodynamic equilibrium for a certain temperature in every respect except in regard to that one process. This really covers the case of a chemical reaction "proceeding at a particular temperature". Strictly speaking, there is no thermodynamic equilibrium, and therefore no temperature, until the reaction has ceased. But all the properties of the reacting substances themselves would be characteristic of the stated temperature, and so the description would be legitimate. In astrophysics we encounter this situation in regard, not to ordinary chemical reactions, but to nuclear reactions.

When we require to distinguish this case from that of strict thermodynamic equilibrium we shall refer to it as *quasi-thermodynamic equilibrium.*

Degree of ionization. Something must be added on this topic in view of its importance in all that follows.

The effect of increasing the temperature of a gas is, by increasing the photon-density and by increasing the violence of the collisions between particles, to increase the efficacy of those processes which remove electrons from atoms. The effect is also, by increasing the speeds of the free electrons and thereby making them harder to recapture, simultaneously

of free electrons which are there to be captured, i.e. to the density of free electrons. Consequently, the temperature of the gas being given, the degree of ionization must increase with *decreasing* density.

Therefore the effect of an increase of temperature upon the degree of ionization is qualitatively similar to the effect of a decrease in density. One important astrophysical application of this is to stellar atmospheres. Two such atmospheres may be in almost the same state of ionization if one possesses a slightly lower temperature, and at the same time a lower density, than the other.

However, the temperature and the density enter into the formulæ in different ways so that their effects cannot completely mask each other.

Again, for the ranges of temperature and density occurring in astrophysical applications, it generally happens that the temperature effect predominates. Thus, for instance, in the interior of a star the temperature is such that the degree of ionization is very high even in regions of great density.

These considerations have been framed with regard being had to the processes producing the ionization. As has been indicated above, in the case of thermodynamic equilibrium the formulæ for the degree of ionization can be derived without reference to these processes. However, such reference not only makes it easier to foresee the nature of the formulæ in this case but is unavoidable in cases to which the formulæ for thermodynamic equilibrium do not apply.

CHAPTER V

SPECTRAL LINES

In a well-constructed detective story there are necessarily three different themes. Though they interact, and to a subordinate degree merge into one another, in the main they are readily distinguishable from each other. Good detective writers are careful to keep the distinction clear. The themes are the facts, the interpretation of the facts, and theories to account for the facts as interpreted. The facts are the existence of the corpse, its condition when found, the presence or absence of finger-prints and so on. The interpretation of these facts give the cause and time of death, the impossibility of suicide and so on. The theories may include one that the murderer was well-known to his victim, that the motive was revenge, and so on until a complete theory of the crime has been constructed.

The way in which these themes react upon one another is obvious. The interpretation of known facts suggests the search for new facts, as for instance the search for the lethal weapon. The theories may also suggest the investigation of facts concerning the activities of certain individuals at certain times, or alternatively may suggest a re-examination of the interpretations. These reactions are what keeps the story going.

The distinction which has been drawn appears to be altogether necessary for making orderly progress with the problems presented. Nevertheless, it is not strictly logical. While it may be represented as a plain "fact" that Smith's dead body has been found in Smith's coal-cellar, all this depends upon interpretation of certain evidence that it is Smith, that he is dead, and so on. This is recognized at the coroner's inquest when these "facts" are established. But the stage at which the court is prepared to accept something as a fact, and not as an interpretation of other facts, though possibly well-defined legally, is quite arbitrary. Then again, the court may be able to accept it as a plain "interpretation"

best made the concerns of different experts: it is the job of the police to amass and classify the "facts"; it is the job of doctors, toxicologists, gunsmiths and so on to "interpret" them; it is the job of the professional detective (or in fiction the exasperating amateur) to produce a "theory" to account for them. It may be added that it is the duty of the jury to produce, if warranted, the conviction.

All this is very like the science of astrophysics. Astrophysics and real-life detection have in common the fundamental feature that their problems are "given" and not manufactured from experiments in a laboratory. But in solving their problems they both, of course, make as much use as possible of laboratory data. Astrophysical problems, too, are solved by the co-operation of the three types of expert we have just described, the observers, the interpreters and the theorists. The same individual sometimes fills all three rôles, but at different stages in his work.

The reason for introducing these considerations at this particular juncture is that while the reader may be prepared to believe in the usefulness of the foregoing chapters for the final theories of our subject, he may have begun to wonder if much of the discussion is at all profitable for the immediate interpretation of the observational data.

Having emphasized the importance of such interpretation, we now recall that in astrophysics a very large proportion of the interpretation is applied to data concerning spectral lines. In the present short chapter we attempt to show how this comes about and how all of the preceding discussion is relevant.

Line profiles. It is somewhat easier to give the following explanation in terms of emission than of absorption; complementary results hold good for absorption and are those most frequently required for applications.

We consider any emission-line spectrum and suppose that it is photographed in the usual way in a spectrograph. Blackening of the photographic plate in the resulting negative denotes the presence of radiation; so the bright lines of the spectrum are portrayed as black lines in the negative. Inspection of the negative shows, however, that "lines" is not a literally accurate description: "streaks" would be a better one. Further, it shows that, though the streaks are all very narrow, they differ from one another both in width and in blackness. Finally, inspection of a single streak shows that it is itself not

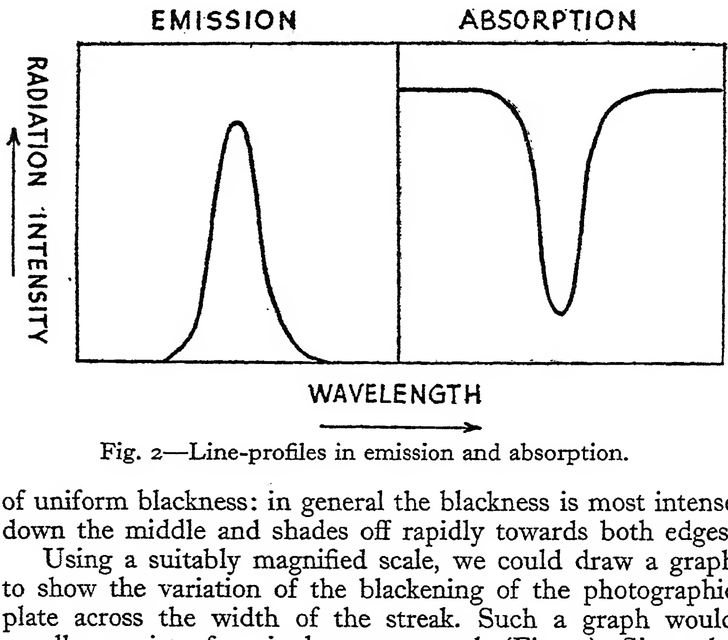


Fig. 2—Line-profiles in emission and absorption.

of uniform blackness: in general the blackness is most intense down the middle and shades off rapidly towards both edges.

Using a suitably magnified scale, we could draw a graph to show the variation of the blackening of the photographic plate across the width of the streak. Such a graph would usually consist of a single narrow peak (Fig. 2). Since the degree of blackening is a measure of the intensity of the radiation falling on the plate, the graph is none other than the energy-curve of the spectrum in the neighbourhood of this particular spectral "line". It is called the *profile* of the line (and we continue to use the conventional term "line" on the understanding that this refers to the whole streak).

An emission line thus consists, not of radiation of precisely one single frequency, but of a concentration of radiation into a narrow range of frequencies. The profile shows the variation of the radiation-intensity over this range.

Line-intensity. When suitably calibrated, the size, that is to say the total area, of the profile is a measure of the total amount of radiation in the line. This we call the line-intensity. We have already seen that this is determined by the number of atoms producing the line.

Shape of profiles. We now consider the shape[1] of the profile. We can indicate the main factors which determine the profile by supposing that we vary the physical state of the gas producing the spectrum.

We first suppose the gas to be disposed in an "optically thin" layer so that a photon emitted anywhere in the gas is practically certain to emerge from it unhindered. We then suppose that we start with the gas at very low density and at almost zero absolute temperature. Assuming that the gas can still emit a detectable amount of the required radiation, we should find for each line a characteristic profile. These profiles would be very much narrower than those normally studied. (Indeed we are here postulating a rather extraordinary instrumental performance in supposing that these profiles could be accurately observed.) The physical process required to explain the profile is called "radiation damping" and is treated by a refinement of the quantum theory beyond the scope of what we have been able to say about the theory. The term is inherited from pre-quantum physics.

Next, suppose we raise the temperature of the gas to, let us say, several thousand degrees. We should find a very marked increase in the width of the profile, particularly noticeable in its central portion. This is called "Doppler widening" or "thermal widening" and is due to the thermal motions of the emitting atoms.

[1] As directly obtained, this is known to be partly determined by the optical properties of the instrument used to give the spectrum. But we are here considering the "corrected" or "true" profile after all merely instrumental effects have been allowed for.

Suppose we then increase the density of the gas without further increase of temperature. When a sufficient density has been reached, we should notice a further increase in the width of the profile, the effect being more pronounced for some lines than for others. The physical process here involved is called "collision-damping" or "phase-disturbance". It is a disturbance of the emitting atoms by collisions[1] with other particles. Under special circumstances, other processes also may yield further "pressure widening" of the profile.

Finally, suppose we take a somewhat thicker layer of gas. Then a photon emitted inside the gas has some chance of being re-absorbed on its way out. Moreover this chance is greatest for frequencies in the most intense part of the spectral line. Thus the profile in each of the foregoing cases must now be less "peaky" than it was with the thin layer of gas. The total intensity of a line is still a measure of the number of atoms producing it, but the measure is no longer one of simple proportionality. The immediate point, however, is that not only the size of the profile, but also its shape must depend upon the number of atoms.

To sum up, the size of a profile depends upon the number of atoms producing it; the shape of the profile depends upon that number, upon the thermal motions of the atoms and upon the pressure in the gas.

Interpretation of profiles. If the state and extent of the gas producing an emission spectrum is known, then, using the principles just enunciated, it is in fact possible to compute the line-profiles. But the problem arising in astrophysics is the converse: from the observed profiles to compute the state and extent of the gas. If this is to be done at all, the method must be in essence to compute the profiles for a set of various assumed conditions of the gas and, by matching the computed and observed profiles, to infer the actual conditions.

Considering the variety of effects which contribute to produce a line-profile, it might appear a hopeless task to attempt to disentangle them by this procedure. In practice, however, advantage is taken of the fact that certain effects

[1] Collisions classed as elastic on page 43.

predominate in certain spectral lines or even in certain parts of their profiles. Also, much may be learned from the comparative study of a number of lines, using for instance the fact that thermal widening depends in a simple manner upon the atomic mass[1] of the emitting atoms. Lastly, there is in general some previously accumulated knowledge of the system under investigation so that some of the factors may be already known. It may therefore be asserted that astrophysicists achieve considerable success in "interpreting" observed line profiles in terms of the physical state of the radiating gas.

We have already mentioned that the case of absorption spectra is more important in practice than that of emission spectra. The atomic processes which affect the profile of an absorption line are the same as those for an emission line. But in general the systems producing absorption spectra are such as to present more difficult mathematical problems regarding the propagation of the radiation through the gas. When these are solved, the problem of interpretation is in principle the same as before.

In subsequent chapters we shall quote information concerning the physical state of the atmospheres of the Sun and stars. A large part of that information is supplied by the methods here described in principle.

Other effects. The preceding discussion has implicitly assumed that the gas responsible for the spectrum is static as a whole and is undisturbed by external forces. A number of further spectroscopic effects can result from various kinds of motion in the gas—streaming, rotation, turbulence, etc.—and from the presence of electromagnetic fields. These effects, too, can in general be disentangled in interpreting the observed spectra, and information can thus be obtained about the motions and forces in the gas.

[1] In thermal motions the mean kinetic *energy* of all types of particle is the same (page 47) so that the mean *speed* is inversely proportional to the square root of the atomic mass.

CHAPTER VI

THE SUN

THE Sun is a typical star—how typical we shall see in a later chapter. Owing to its obvious dominance of terrestrial affairs, man would have studied the Sun had there not been another star in the sky. But to man as an astrophysicist the Sun is a typical star that can be studied with incomparably more accuracy and detail than any other of the myriads of stars which it typifies. This is its significance to the astrophysicist and it is from this standpoint that the present chapter attempts to discuss our knowledge of the Sun.

General data. We start with a brief survey of the basic observational data[1] concerning the body we have to study.

The Sun is a sphere of diameter 1·4 million kilometres and of mean density nearly one-and-a-half times the density of water. It rotates about its axis, though not as a rigid body, with a period of 27 days in mean latitudes.

The Sun radiates at the rate of about 8 horse-power per square centimetre of its surface. This is the rate of radiation from a "black-body" at nearly 6,000 degrees. There is evidence from the history of the Earth's surface to show that this rate must have been approximately constant for at least several hundred million years.

There is a good deal of evidence—but it is not conclusive—that the Sun possesses a general magnetic field like the general magnetic field of the Earth. The ascribed field intensity is, however, too small to have any significant effect upon other general characteristics of the Sun.

When the Sun is viewed directly its disk appears to have a perfectly sharp edge. But when this disk is obliterated by the Moon during a total eclipse, the Sun is seen to be surrounded

[1] They are put in round numbers in the text: more exact figures are quoted in Table I.

r an "outer atmosphere" consisting of a layer about 10,000 lometres thick called the *chromosphere*, and a much more nuous envelope having a depth of the order of a million lometres called the *corona*. We return later to a brief study this outer atmosphere; meanwhile we shall treat the Sun though shorn of such an appendage. This is not to imply e existence of a perfectly definite interface between the Sun ıd its chromosphere, but the transition layer is known to be latively very thin. The outermost layer of the Sun proper, hose outside boundary provides the apparent "surface" of e Sun, is called the *photosphere*; it will be more fully described due course.

TABLE I

THE SUN

Mass	$1{\cdot}98 \times 10^{33}$ grammes
Radius . . .	$6{\cdot}95 \times 10^{10}$ centimetres
Luminosity . . .	$3{\cdot}72 \times 10^{33}$ ergs per second
Spectral class . .	dG2
Rotation-period . .	27·5 days in latitude 40°

Mean density	1·41 grammes per cubic centimetre
Radiation flux from surface .	$6{\cdot}13 \times 10^{10}$ ergs per square centimetre per second
Effective temperature . .	5,713 degrees
Mean radiation per unit mass	1·88 ergs per gramme per second

The data for the Sun are represented in Figures 3–8 ɔy the symbol ⊙.

lar spectrum: Fraunhofer lines. The spectrum of solar liation consists of a continuous background interrupted by :k lines. Such lines were first observed in 1802 by W. H. ɔllaston and were first extensively studied in 1814 by

J. Fraunhofer after whom they are called. The first ide tion of a chemical element in the Sun, by the aid Fraunhofer lines, was that of sodium by G. Kirchhoff in In 1872 H. Draper was the first to publish a photog spectrum of the Sun and so to inaugurate the modern of its accurate analysis. In 1893–6 H. Rowland publis great "map" of the spectrum together with a table of lengths of some 20,000 Fraunhofer lines; many more since been added to the list.

By means of these wave-lengths it has been possi assign most of the strong Fraunhofer lines and thou of weaker ones to chemical elements known on the No one has any doubt that the remaining lines could a so assigned.

Up till now, sixty-odd of the 90 terrestrially known ele have been identified in the Sun, a few of them doub and a few not in the Fraunhofer spectrum but only in s of prominences, etc. (see below). It is known that ele not so identified, if present in likely quantities, wou exceedingly difficult to detect owing to the particular p of their spectral lines.

The vast majority of the lines belong to the ator ionized atoms of the elements concerned, but the solar spe contains also many lines due to molecular combinatio atoms (such as cyanogen). Except as useful suppleme indicators of physical conditions, the presence of such pounds is unimportant for present purposes and ca ignored; it is known that they can exist only in the oute fringe of the Sun's photosphere.

Almost from the outset of the study of Fraunhofer their differing widths and darknesses did not pass unno The lines in Kirchhoff's catalogue of 1860 were clas according to these characteristics. Rowland's table assign each line an estimated "intensity". His scale has subsequ been calibrated so as to bring it into as close accord as pos with the modern measure of intensity and it continues to source of valuable information. The first measurement actual profile of a Fraunhofer line was made by K. Schw child in 1914 and the first applications of the metho

photographic photometry used ever since were made in 1927 by several workers in Potsdam and Utrecht.

The culmination of such work in regard to general surveys of the solar spectrum has been the publication in 1940 by M. G. J. Minnaert, J. Houtgast, and G. F. W. Mulders of their monumental *Photometric Atlas of the Solar Spectrum.* This is no less than a plot, on the extremely accurate scale used in the work, of all profiles recognizable as such in almost the whole range of wavelengths accessible to photographic photometry. In other words, it constitutes an exceedingly detailed energy-curve of the solar spectrum, showing the details accurately in relation to the neighbouring background but not, what is irrelevant in this context, the *general* variation of the background.

Solar spectrum: continuous. The general variation of the background, that is to say, the nature of the continuous spectrum of the Sun, presents a very different problem. It has been the subject of much careful study both from the point of view of the shape of the energy-curve and of absolute light-intensity.

In general appearance, the energy-curve[1] of the continuous spectrum is similar to a Planck curve for about 6,000 degrees but has a narrower peak and falls below the Planck curve in the ultra-violet.

We have already quoted a rough figure for the total radiation of the Sun. Accurate measurements of absolute intensity are exceedingly difficult on account of the corrections required for the loss of radiation in the Earth's atmosphere and in the instruments, and also on account of the difficulty of establishing absolute standards for comparison. The most extensive series of observations have been made by C. G. Abbot, but many others have contributed to our present knowledge. A. Unsöld's discussion of the data gives the figure quoted in Table I.

If the Sun were replaced by a black-body of exactly the same size this would give the same total radiation if the black-

[1] The energy-curve of the continuous spectrum can mean rather different things according as we suppose the Fraunhofer lines to be *filled in* or to be *ironed out*. We can afford to ignore this distinction in the present discussion.

body temperature were 5,713 degrees. This is called the *effective temperature* of the Sun. Owing to observational uncertainties in the total radiation, this value may well be in error by about 30 degrees either way.

It has to be emphasized that, although we speak of the effective temperature "of the Sun", it is in fact the *temperature* of a hypothetical black-body giving the same total radiation. So far as the Sun itself is concerned, it is only a convenient measure of this total radiation. It begs no questions about the actual temperature of any part of the Sun.

Abbott has found evidence of fluctuations of the order of a few per cent in the total radiation of the Sun, such fluctuations occurring in anything from a few days to a few years. They affect the ultra-violet end of the spectrum and are almost certainly associated with solar activity (see below). Neither these nor any other observations provide any evidence that the Sun is to be classified as a "variable" star in the ordinary sense of this term.

Thus far we have been speaking of *integrated radiation* from the Sun, which is the mixture of radiation from all parts of the solar disk "just as it comes". This is all that we can normally hope to study in the case of any other star. But in the case of the Sun it is possible to cause a telescope to give an enlarged image of the disk, and then to admit to the spectrograph the radiation from only a very small portion of the image. In this way it is possible to study the variation of the radiation from the centre to the limb[1] of the disk, which is the same thing as studying its variation with direction of emission from any one locality of the spherical surface of the Sun. Or, again, it is possible to study the spectra of spots and other markings on the surface and also, particularly during an eclipse, to go outside the disk to obtain spectra of various features of the Sun's outer atmosphere.

Photosphere

We call the outer part of the Sun, down to a certain depth (to be specified later) below anything that we can see directly,

[1] "Limb" is always used in astronomy to mean "edge", i.e. the word has the Latin and not the English root.

the kind observed serves to establish its *gaseous* state. Indeed, without making any measurements, a spectroscopist infers from the wealth of well-defined lines in the Fraunhofer spectrum that it is produced by a gas which, as judged by laboratory standards, is at a high temperature and a low pressure. The fact shows also that we are seeing through some of this gas, not just looking at a quite definite surface of the Sun. For the essence of an absorption spectrum is the absorption by certain material of light *traversing* that material.

Next, the fact that energy is streaming away from the Sun is sufficient to demonstrate on general thermodynamic grounds that *the temperature must increase inwards* from the boundary. This is shown also by the mere existence of the Fraunhofer lines. For, were the temperature uniform through the photosphere right up to its boundary, the spectrum would be purely that of a black-body and would contain no lines at all. Therefore, in viewing the Sun we are seeing into layers not all at the same temperature. And, of course, this variation of temperature cannot be a decrease inwards since this would require a resultant flow of energy inwards.

[In view of the explanation given in Chapter IV, all we can strictly conclude from the argument just stated is that the photosphere as a whole is not in thermodynamic equilibrium. However, the general practice in astrophysics, save when there are manifest reasons for the contrary, is to consider first in such problems the consequence of postulating *local* thermodynamic equilibrium. If the consequences agree with

[1] The study of the outer atmosphere does show that this region has very little effect upon the normal solar spectrum. If, however, the reader prefers to suspend judgment on this point and to take the photosphere provisionally to include any part of the outer atmosphere which, until the contrary is demonstrated, could appreciably affect this spectrum, then the ensuing arguments still hold good.

Some readers may remark upon the apparent neglect of what used to be called the "reversing layer"; this is explained later.

observation, there is no need to change the postulate; if not, it becomes necessary in the light of the disagreement to see how the postulate has to be modified. By speaking above in terms of temperature we have already implicitly made the postulate in the present case. It must be stated that it is in fact found to give a good first approximation when the quantitative results are worked out and compared with observation.]

Any photon which enters an observer's eye or his instruments from the Sun must have had a definite point of departure *P* in the photosphere. At *P* it was either emitted by an atom or was deflected, for the last time, by a material particle in the photosphere. Another way of saying that we see into the photosphere is to say that the points like *P* are distributed throughout some depth of the photosphere. Moreover, the opaqueness of the photosphere is bound to be different for radiation of different frequencies. So photons of frequencies for which the opaqueness is small come from a greater average depth than those for which it is great. This is the phenomenon taken into account in "infra-red" photography of terrestrial landscapes: atmospheric haze is less opaque to infra-red than to higher frequencies.

An immediate consequence of these various considerations is that the energy-curve of the Sun's continuous spectrum cannot be that of temperature-radiation for a single temperature, but must be, in effect, a blend of such radiation for the range of temperatures corresponding to the range of depth to which we see into the photosphere. Also the composition of the blend must be expected to be different for different frequencies.

Another consequence is that within the narrow range of frequencies covered by an absorption line we are seeing down to a much shallower average depth of the photosphere than in frequencies in the continuous spectrum.

This was formerly expressed by distinguishing two layers in the solar atmosphere: the "photosphere", from which the continuous radiation was considered mainly to come, and, overlying the photosphere, the "reversing layer", in which the absorption lines were considered mainly to be formed. This

distinction has now been largely discarded: it is too indefinite because, even for a single line, different parts are formed at different average depths, and also because there is no essential distinction in the general physical conditions such as that which may be drawn between the photosphere and the "chromosphere" or that between the photosphere and the "interior" of the Sun.

Finally, if we look at a point of the solar disk near the limb we see down to a smaller average depth than if we look at a point near the centre of the disk. For, owing to the curvature of the photosphere, we are looking into it obliquely near the limb and so we encounter the same amount of opacity in penetrating to a smaller radial depth as compared with what happens when we look vertically into the photosphere near the centre of the disk. Hence, near the limb, we are seeing into a layer of gas of more nearly uniform temperature than near the centre. We should therefore on this ground expect the Fraunhofer lines to be relatively less intense there than near the centre. Also, we are looking into gas with a lower average temperature than when we look at the centre and so its intrinsic brightness will be less. The disk must therefore exhibit the phenomenon of "darkening towards the limb". Thus the Sun must actually appear globular; this would not be so were it to behave as a black-body when it would present a disk of perfectly uniform brightness.

It may be asserted that the qualitative inferences of this section are justified by the appropriate comparisons with observation and by more detailed investigation. The only exception is in regard to the centre-limb variation of the Fraunhofer lines: for the reasons stated there is such a variation but it is more complex than we have suggested. The main point of this section has been to show that we *can study the photosphere in depth*, that we can, in fact, though perhaps only to a puny extent, delve into the Sun and, literally, see what is there and what it is doing there.

Energy transport. The total radiation of the Sun constitutes a steady flux of energy from its "surface". Now even the most rudimentary estimate of the physical state of the photosphere

disposes of any suggestion that this energy is being generated by chemical, sub-atomic or any other processes taking place in the photosphere itself. Consequently, this energy must be *transported* through the photosphere from the interior to the surface: the question is, how?

Energy can be conveyed from one place P to another place Q only by matter or by radiation. If the transport is by matter, it may be achieved by the particles of the matter passing their energy from one to another by processes essentially similar to atomic collisions, without any bodily motion of the matter as a whole: this is transport by *conduction*. Or else matter endowed with energy may move bodily from P to Q (where, if necessary, it may liberate some of its energy and then return to P): this is transport by *convection*. In the case of transport by *radiation*, if there is matter intervening between P and Q, photons may be absorbed and others re-emitted on the way, or photons may be scattered by the matter, but all the "travelling" energy is in the form of radiation.

All three means of transport are illustrated by a steam radiator. Heat is convected from the furnace to the inside of the radiator, then conducted through the metal, and finally radiated from the surface.

In the solar photosphere the transport is through a gas at low pressure. The known order of magnitude of the thermal conductivities of gases shows that no significant part of the observed energy-flux could be conveyed by conduction. We have therefore to decide between the rival claims of convection and radiation.

If we have any medium through which there is any energy-flux, we know on general thermodynamic grounds that there must be a variation of temperature in the medium and that the flux is from the hotter to the cooler regions.[1] But we also know on very general grounds that, in any medium which is not completely opaque to all radiation, there must be a net flux of radiation from the hotter to the cooler regions. Combining these two conclusions, we see that if there is any flow

[1] This has been stated previously and it has been explained that the statement may be expressed in terms of temperature only if we assume local thermodynamic equilibrium.

case, it is possible to calculate whether they will be set up or not.

Consider then the problem of the photosphere. The astrophysicist knows that it is a gas maintained in a steady state under solar gravitation with a known energy-flux through it. He first makes the trial hypothesis that there is no convection. He can then calculate the state of the gas, and in particular the temperature distribution through it. He then examines whether these calculated conditions would or would not result in convection being set up. If they would not, then this trial hypothesis is justified and he concludes that the photosphere is in *radiative equilibrium* and that, necessarily, the entire energy-flux is radiative. If the calculated conditions would result in convection, then his trial hypothesis that there is no convection has produced an inconsistency. Hence there must be convection and the photosphere must be in what is called *convective equilibrium*.

In the latter case the astrophysicist has to repeat all his calculations, as he perfectly well can do, on this fresh assumption and he thus derives the actual state of the gas. But he knows that, as we have said, there must still be some radiative flow of energy. So he finally has to calculate its amount in this state of the gas. He may find that, in spite of the presence of convection, almost all the flux is radiative, in which case the photosphere would be in convective equilibrium but with radiative energy transport. Or he may find that the convection accounts for almost all the transport, so that the photosphere would be in convective equilibrium with convective transport.

What is actually found in the case of the photosphere will be noted below. The reason we have described this procedure rather fully is that it can be followed equally well in the study of the interior of the Sun and leads to very important conclusions in that context.

re-emitted in the same frequency, the re-emitted radiation escapes in all directions and very little of it re-enters the direct beam. Therefore the direct beam is certainly rendered deficient in radiation of this particular frequency and so, when analysed spectroscopically, it exhibits an absorption line at that frequency.

The case of the photosphere is different in two regards: (a) any portion of material within the photosphere is illuminated from all sides and not by a "beam" of light, though there is a net flux of radiation in one direction; (b) the radiation which such a portion emits does not escape directly from the system but, in fact, helps to provide the radiation illuminating other portions. It must suffice to say that this situation can be handled mathematically and the intensity and profile of an absorption line may be related to the properties of the photospheric gas according to the principles summarized on page 56.

Photosphere—constitution. We must now quote results with only scanty further indications of their derivation. It is hoped that the reader has been put in possession of the underlying physical principles, and of the main lines of the problems to which they are applied, to an extent sufficient to enable him to see that the derivation is actually possible. Also, he will readily appreciate that the results are the fruits of cross-fertilization between the study of the continuous radiation and that of the line-spectrum. The following paragraphs summarize the emergent picture of the photosphere.

The photosphere is composed almost entirely of *hydrogen* gas. The other constituents which are significant for its general behaviour are *metals*—in gaseous form. There is about one metal atom to every 10,000 hydrogen atoms. The hydrogen is

effectively not ionized while the metals are nearly completely (singly) ionized. Thus there is about one free election to every 10,000 hydrogen atoms; as regards the general physical conditions, the metals are important mainly as the suppliers of these free electrons.[1]

Those atoms of any element which contribute directly and appreciably to the formation of a particular Fraunhofer ine amount at most to about one-thousandth of a gramme per square centimetre of the Sun's surface. They may amount to vastly less; for instance, the hydrogen lines are ascribed to about one-hundred millionth of a gramme per square centimetre of the surface (composed of atoms in the quantum state producing the lines). Thus an element may be present in sufficient quantity to contribute to the Fraunhofer spectrum, but in a sufficiently small *proportion* to have no appreciable effect on the general physical conditions.

Most of the radiation emitted by the Sun comes from the irst few hundred kilometres below the outer surface of the photosphere—about 300 km. would be a fair estimate to keep n mind. In the terms used before, the majority of photons which escape from the Sun have their points of departure in his layer. The layer has a mass of only some 3 to 5 grammes per square centimetre of the surface. The temperature at the urface is about 4,800 degrees and the average temperature hrough the layer is about 5,700 degrees. The total gas pressure at the bottom of the layer is of the order one-tenth of an atmosphere and the pressure of free electrons about 10 nillionths of an atmosphere. In order to give more precise esults it would be necessary to go into much more detail regarding the variation of temperature and pressure with depth.

The most astonishing feature revealed by these figures is he great opacity of the gas in regard to the continuous radiaion, that is, in effect, to ordinary sunlight. Whereas the Earth's atmosphere with over 1,000 grammes to the square entimetre is highly transparent to this radiation (in the main part of the frequency range) a layer of the Sun's atmosphere

[1] Certain other elements, notably helium, are present in greater bundance than the metals. But they are not important as a source of ree electrons in the photosphere.

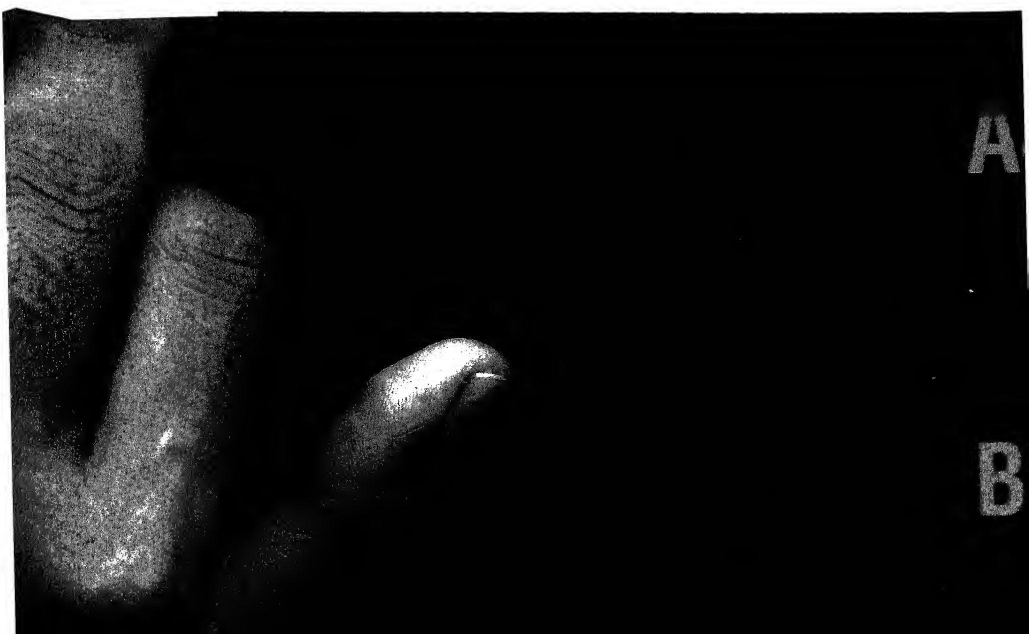
others with the result that the energy curve would be grossly different from anything like a black-body curve. As we have seen, however, the over-all departure from a black-body curve is not excessive.

Continuous opacity. It was not until 1939 that an apparently satisfactory explanation of this opacity was forthcoming when R. Wildt advanced the view that it is produced mainly by *negative hydrogen ions*. This is an instance of an important astrophysical problem demanding for its solution an atomic phenomenon which plays no important part in terrestrial physics and the understanding of which depends upon a considerable refinement of atomic theory.

Other negative ions are known and are of some astrophysical interest, but their importance is not comparable with the case of hydrogen to which we confine discussion.

Although the normal hydrogen atom consisting of the hydrogen nucleus, or proton, and one electron is of necessity electrically neutral as a whole, the electron cannot completely shield the electric field of the proton at very short distances. So it is not surprising to find that the neutral atom can in fact accommodate a second electron and thus become, as a whole, an "ion" with a net negative charge of one unit. This combination has been thoroughly studied by the aid of quantum theory. As we should expect, the additional electron is very loosely held: the energy required to remove it is only about one-twentieth that required to remove the original electron from the normal atom. It has no excited states and so yields no line-spectrum. But its bound-free transititions render it an efficient absorber of continuous radiation in the frequency range of the intense part of the solar spectrum.

In order that a negative ion should be readily formed in a gas it is no use to expect the extra electron to be provided by



another hydrogen atom. For any agency that is sufficiently energetic to remove the electron from the other hydrogen atom would *a fortiori* remove it from the desired negative ion. In fact, any suggestion that negative ions might be formed in pure hydrogen calls to mind the familiar jokes about Aberdonians collecting from Aberdonians. The Aberdonian (according to the convention of the stories) must always collect his subscription from someone from whom it is more easily detached than from himself. Similarly, the negative ion must collect its extra electron from some atom from which it is more easily detached than from another hydrogen atom.

Now the photosphere furnishes precisely the conditions required for a successful collection: there is abundant hydrogen, all un-ionized, together with a supply of free electrons provided by the much more easily ionized[1] metals. Thus negative ions can be formed and their absorptive power brought into operation. Calculations made by several astrophysicists show in fact that the admixture of hydrogen and metals in the proportions found in the photosphere does endow the photosphere with the astonishing opacity which it is found to possess.

To sum up, the opacity of the photosphere is held to be produced by electrons performing bound-free transitions in the fields of neutral hydrogen atoms, these electrons being supplied by the ionization of metal atoms. It is a remarkable result that, whereas pure hydrogen in the conditions here considered would be effectively quite transparent, the introduction of one metal atom for every 10,000 hydrogen atoms renders it highly opaque.

The idea of an opaque gas is, however, in itself nothing novel. Although, as we have said, the Earth's atmosphere is highly transparent to ordinary sunlight, it is completely opaque to ultra-violet radiation from the Sun. As is well-known, this opacity is due to ozone in the Earth's upper atmosphere, and we might almost venture (chemistry apart!)

[1] "Ionized" and "un-ionized" here and always refer to the *removal* of electrons; if an electron is removed from a neutral hydrogen atom so that a positive ion is formed we say that the atom is ionized; if an electron is added to a neutral atom we do not say that it is ionized even though the product is called an ion (a negative ion in this case).

to call the negative hydrogen ions in the photosphere the Sun's "ozone layer".

In the case of the Earth's atmosphere the radiation comes from outside and the absorbing layer stops the radiation (in the appropriate frequencies) from getting through. In the case of the photosphere, the radiation comes from inside; the continuous absorption does not stop the radiation, but it causes the photons to be many times absorbed and re-emitted in their passage through the photosphere with the result that we can see into only a very small mass of the gas.

The calculations show also that the dependence of the resulting opacity upon frequency is such as to account for the form of the main part of the energy-curve. In the ultra-violet and in the far infra-red regions the continuous opacity is produced mainly by effects other than those due to negative hydrogen ions. These effects are well-understood and are more commonplace than the latter. But we cannot discuss them here beyond mentioning that they must produce the whole of the continuous opacity at a sufficient depth in the photosphere, since the negative ions cannot exist to any appreciable extent in regions where the temperature exceeds about 6,000 degrees.

Convective layer. At any level in the photosphere, the pressure is that required to support the weight of material above that level. Hence the pressure increases with increasing depth. Also we have seen that the temperature must also increase in order to maintain the outward flux of energy. The material being gaseous, the increasing pressure is produced partly by this increasing temperature and partly by increasing density. But increasing density means increasing opacity, and this requires an increasing rate of increase of temperature in order to force the necessary energy-flux through the material. We see, therefore, that the temperature must increase more and more steeply with depth as the depth increases. All this can be translated into mathematical terms and the temperature, pressure, etc., can be evaluated at all depths.

At the surface of the sun all the radiation is going outwards and so the net outward flux represents all the radiation present

at that level. But at any deeper level there is radiation travelling in all directions, and this adds up to give approximately the same density of radiation as would be present in an enclosure in thermodynamic equilibrium at the temperature for this depth (or exactly this density, if we use it to define the temperature as it is in fact convenient to do). Since then the temperature increases inwards, so also does the density of radiation. The net outward flux of radiation, on the other hand, is the same at all levels in the layers we are considering. Therefore the net outward flux represents a smaller and smaller fraction of the total radiation as the depth increases.

This enables us to complete our definition of the photosphere. We can say that it includes all levels at which the net outward flux is a significant fraction of the total radiation. We could take a significant fraction to mean, say, one per cent or more: it does not matter much what we take since the temperature is increasing so rapidly at the base of the photosphere that any other reasonable value of the fraction would place the base at approximately the same level.

Now we have stated that the hydrogen in the photosphere is virtually un-ionized. Owing, however, to the rapid increase of temperature with depth in the neighbourhood of the base of the photosphere, appreciable ionization does set in there. So the photosphere merges into another layer in which the degree of ionization of hydrogen increases from practically zero to practically unity. Below this layer, that is in fact throughout almost the whole mass of the Sun, the hydrogen is completely ionized.

The changing degree of ionization in this transition zone has a most interesting consequence. Suppose a portion of the gas at any level in the layer is brought to a higher level and there allowed to attain the pressure of its new surroundings. The drop in pressure produces a drop in temperature. But a drop in temperature results in a drop in ionization. This in turn means the release of some of the original energy of ionization and this released energy tends to heat the gas up again. If now the temperatures at the various levels were those calculated on the hypothesis of radiative equilibrium, then it works out that the net result of the release of ionization-energy in the

transported gas would be to leave it hotter than its new surroundings. Consequently this portion of gas would go on rising just as hot air rises to the top of a building. Similarly, had the portion of gas been brought to a lower level it would go on sinking. Thus the whole layer would be in an unstable state. This is the consequence of supposing it to be in radiative equilibrium.

We conclude therefore that the layer is in convective equilibrium and term it the *convective layer*. It is estimated to have a thickness of the order of 500 km. Though its equilibrium is convective, energy-transport through the zone is still mainly radiative, probably only about 10 per cent of the energy-flux being carried by convection.

Granulation. The surface of the Sun is not of absolutely even brightness: it exhibits a fine-grained appearance of rapidly fluctuating detail. The granules have a mean diameter of the order of 1,000 km. and a mean life of about 2 minutes. They are explained as being "convection cells" which tend to be set up in a layer of fluid heated from below when it cannot transmit the heat sufficiently quickly by other means. Under suitable conditions, such cells are produced in the Earth's atmosphere and are evidenced by the appearance of a "mackerel sky".

It is considered that convection cells must be set up in the convective zone. Since the photosphere is a layer of relatively small mass floating on top of this zone, it is considered further that the cells must penetrate the photosphere and so produce the observable granulation of the Sun's surface. Thus we appear to have fairly direct evidence of the existence of the theoretically inferred convective zone.

Solar activity and the solar cycle. The existence of spots on the Sun has been known for over 300 years. In quite recent times it has become known to most people that these spots are only one manifestation of a group of phenomena called *solar activity* which owes its notoriety to the fact that it can so seriously interfere with radio reception.

Solar activity is a fascinating subject. Its various manifestations are probably the most intensively studied phenomena

10 light upon any general problem of astrophysics. It would herefore be inappropriate to devote much space to it here. Nevertheless we must give some brief description of the phenomena themselves, firstly in order to show that the Sun's atmosphere is not such an uneventful region as might be supposed from the rest of our study, and secondly in order o indicate what general problems are probably connected vith the phenomena.

It is well-known that the number of spots on the Sun follows fairly regularly an 11-year cycle from minimum spottedness through maximum spottedness and back again to ninimum. At minimum there are scarcely any spots; then a few start to appear in solar latitudes about $\pm 30°$; as the number as well as the average size of spots increase, they occur in progressively lower latitudes until at maximum their nain occurrence is spread over two belts each of width about 20° of latitude separated by a narrow equatorial strip; finally the numbers occurring in these belts diminish until there are very few left, when the cycle starts afresh; the interval from ninimum to maximum averages about $6\frac{1}{2}$ years. The details of these proceedings are subject to erratic variations.

An individual sunspot appears as a relatively dark region on the Sun's disk, the central portion or *umbra* being darker than the surrounding *penumbra* with a fairly well-defined demarcation between them. The total area may be anything up to nearly one-half per cent of the Sun's visible hemisphere. The majority of spots last only a few days, but a minority of spots or groups of spots persist for weeks or even months.

A sunspot looks dark only by contrast with the surrounding photosphere. Intrinsically it is quite bright: it is in fact just an area of the Sun's surface some 1,500 to 2,000 degrees cooler (in the umbra) than the normal photosphere.

A spot is found spectroscopically to have associated with it a strong local magnetic field whose lines of force are radial

to the Sun near the centre of the spot, i.e. the spot-field is like that of a bar-magnet floating vertically in the photosphere. The pole which is uppermost may be "north" or "south" according to certain empirical laws. During a single 11-year cycle, the laws which hold for the Sun's northern hemisphere are the opposite of those which hold for the southern hemisphere. But the most remarkable fact is that the laws which hold for one cycle are the opposite of those which hold for the next cycle. When the magnetic characteristics of the spots are taken into account the complete period of solar activity between successive repetitions of the same phenomena is therefore not eleven but twenty-two years.

The regularity, in their broadest aspects, of these occurrences is thought to show that their ultimate cause is something deep-seated in the Sun. This cause, which may arise from some circulation of material inside the Sun, would impose the 11- or 22-year periodicity. The equatorwards progression of the manifestations of activity is then likely to be due to the fact that the Sun does not rotate as a rigid body: its surface rotates more quickly near the equator than near the poles.

The darkness of a sunspot is easily interpreted as being due to the cooling of material brought up to the surface from below and there expanding under the reduced pressure of its new surroundings. This interpretation is substantiated by the fact that the gases of the Sun's atmosphere above a spot are seen spectroscopically to be flowing outwards over the surface from the spot with speeds of a few kilometres a second ("Evershed effect").

The origin of the magnetic field of a sunspot is completely unknown. Opinions differ as to whether the magnetic field (cause unknown) produces the outflow of material which constitutes the spot, or whether the outflow (cause unknown) produces the magnetic field.

Thus the broader aspects of solar activity are expected ultimately to lead to increased knowledge of solar hydrodynamics while the more detailed aspects are likely to yield increased knowledge of the structure of the Sun's atmosphere and also of the origin of magnetic fields in astronomical bodies.

all the phenomena shares in the periodicity of the 11-year le.

It must first be emphasized that the photosphere as a whole, :luding those regions immediately affected by some form of ivity at any particular moment, shows no sign whatever of y general periodic variation in any of its features. If then ar activity does in fact spring from some cause deep within e Sun it must be something which transmits its influence the surface by some very strictly localized agency.

The phenomena to be mentioned are the following:

lar flares are sudden short-lived intense brightenings of nall regions of the Sun's surface usually in close proximity sunspots. They are known to be the source of ultra-violet diation causing radio "fade-outs" by its effect upon the arth's ionosphere. R. G. Giovanelli has proposed a theory flares according to which the growth or decay of the magetic field of a sunspot produces an electric current around inder and over) the spot and, by a combination of the local iagnetic field and the general field of the Sun, this current roduces intense heating in a particular region. This region cts rather like the filament of an electric lamp in an electric ircuit and so yields the flare.

rominences are clouds of luminous gas, tens or hundreds of housands of kilometres in extent, projecting from the solar urface or suspended above it. They may be quiescent for lays or may exhibit complex eruptive changes in a matter of ninutes. They are not confined to the sunspot zone. Prominencies rank as the most spectacular occurrences in the solar atmosphere but they have so far defied explanation, at any rate in terms acceptable to the general body of solar physicists.

Faculæ are certain markings on the Sun's surface revealed by suitable instruments. Prominences as such are readily observed only when they are seen in elevation at the Sun's limb; faculæ are almost certainly prominences when seen in projection on the Sun's disk.

Eruptions also apparently occur with the emission of streams of ionized gas which provides the so-called "corpuscular radiation" causing magnetic storms and auroræ when it reaches the vicinity of the Earth. The gas traverses the intervening space with a speed of about 1,000 km./sec. and its ejection from the Sun is generally ascribed to the operation of radiation pressure. Such eruptions can frequently be associated with regions of the Sun showing other forms of disturbance, but they have not been uniquely identified with any particular form of visible disturbance.

Solar noise shows violent fluctuations associated with spots and flares, but we shall return to this when we mention the subject as a whole.

Outer Atmosphere

Chromosphere. Immediately above what we have been treating as the surface of the Sun there lies a tenuous layer of gas about 14,000 km. thick called the *chromosphere*. It is so tenuous that the continuous radiation from the photosphere traverses it quite unhindered and it has only small effects upon the Fraunhofer lines. But during a total eclipse of the Sun it can be seen at the edge of the disk without the background of direct sunlight, and, as is to be expected, it then gives an *emission* spectrum. This is the famous "flash" spectrum, so called because it flashes into visibility at the instant when the eclipse becomes total; its detailed study yields knowledge of the distribution of density in the chromosphere.

The puzzle of this density distribution is that it appears at first sight to defy solar gravity; the density is found to decrease with height far more slowly than it would do if the gas is in hydrostatic equilibrium and if its temperature is about equal to the Sun's surface temperature, i.e. about 5,000

degrees. The immediate explanation is that the gas is agitated by velocities considerably in excess of those corresponding to thermal agitation at this temperature. The existence of such velocities is shown by spectroscopic evidence and, in a general way, by the fact that observers describe the detailed appearance of the chromosphere as that of a "billowing meadow" or "burning prairie". There is no generally accepted explanation of these velocities, though the very recent theory of the corona by H. Bondi, F. Hoyle, and R. A. Lyttleton seems to account for them as a necessary feature of the transition region between the corona and the normal atmosphere.

Corona. If the chromosphere gives the impression of defying the law of gravitation, the solar corona not only does this on a much more magnificent scale, but appears to defy the laws of thermodynamics as well! It is a rarefied atmosphere around the Sun extending out to distances greater than the solar radius and sometimes exhibiting "streamers" traceable to many times this distance. During a total eclipse it gives about half as much light as the Moon when full, or about one-millionth the light of the Sun. Only the most refined technique enables it to be detected except during such an eclipse.

The light of the corona is partly scattered continuous sunlight and partly that of the line emission spectrum of the corona itself. There are some twenty lines of this spectrum, but not one of them has even been observed in any laboratory spectrum. Nevertheless, by one of the most notable triumphs of modern spectroscopy, the majority of the lines have recently been identified by B. Edlén as arising from calcium atoms which have lost 11 or 12 atoms, iron atoms which have lost 9–13 electrons, and nickel atoms which have lost 11–15. The identifications depend largely upon the extrapolations of known empirical regularities concerning quantum levels and transition [illegible]abilities. The internal consistencies in the results have l[illegible]ir general acceptance despite the alarming impli[illegible]e "temperature" of the relevant regions [illegible] the [illegible] of the order of a million degrees in o[illegible] for such [illegible]ation to occur. At any rate, wh[illegible]

agency produces the ionization must be characteristic of such a temperature.

The existence of a "temperature" so hugely in excess of that of the underlying solar surface does not actually contradict any fundamental thermodynamic principle when we take account of the transparency of the corona and the very small amount of matter concerned. The total mass of the corona is probably only a few millionths of a gramme per square centimetre of the Sun's surface. The difficulty is to explain the presence of the corona and the cause of its heating: whatever "temperature" results from this heating is purely the affair of the corona itself—it cannot have any appreciable effect upon the thermodynamics of the Sun as a whole.

Innumerable theories have been proposed and discarded. At the present time there is however a growing measure of support for those which treat the coronal gas as being supplied by material falling into the Sun. Such material might be either interstellar (Chapter I) or interplanetary (i.e. meteoritic, though such matter might itself first have been collected into the solar system from interstellar space). In either case the heating is ascribed to the conversion of the kinetic energy acquired by falling under solar gravity. This energy would in fact be adequate to produce a temperature of the required order. When the presence of the material and its temperature are accounted for in this way, they, in turn, account satisfactorily for the density distribution in the corona. Bondi, Hoyle, and Lyttleton have shown that the theoretical consequences of supposing the incoming material to be largely interstellar hydrogen reproduce the main observed characteristics of the Sun's outer atmosphere.

The corona possesses a complicated structure composed of "rays" and "streamers"; at times of maximum solar activity it is usually fairly evenly distributed round the Sun, but at minimum it tends to be relatively less extended along the solar axis and more extended near the equatorial plane. These various features are usually regarded as not being inherent in the production of the corona but as being impressed upon its highly ionized material by electromagnetic effects associated with solar activity.

Solar Radio Noise

Undisturbed Sun. We have seen that in visible light the Sun radiates approximately like a black body at a temperature of about 6,000 degrees. Were this true for radiation of all wavelengths, then, by simply drawing the rest of the black body curve, we should have an estimate of the intensity of the radiation emitted by the Sun in any wavelengths outside the visible range. In particular, we should in this way predict that the Sun must radiate a certain amount of energy in the wavelengths of radio waves. But the amount so predicted is very small and there would be little hope of detecting it with any certainty.

The astonishing fact is that, as extensive observations made in recent years have established, the Sun does emit readily detectable amounts of radiation in radio wavelengths between a few metres and a few centimetres. The intensity reaches values about a hundred thousand times greater than the emission corresponding to a temperature of 6,000 degrees. The radiation has become known as *solar radio noise.*

The most intense emissions are associated with the occurrence of spots and flares. The "quiet" Sun, i.e. the Sun when free from such disturbances, has however been shown to maintain a steady noise output. This is greatest on about 1·5 m. wavelength and its intensity at that wavelength is not that corresponding to a solar temperature of 6,000 degrees but to one of about a million degrees.

Startling at these phenomena may appear, they do reconcile themselves in a most satisfactory fashion with the view of the Sun's outer atmosphere which has been sketched in the preceding section and so furnish a remarkable confirmation of that view.

We learned that the corona is a gas in an extremely highly ionized condition. Now we know that an ionized gas can be highly opaque to radiation of radio wavelengths. Much of our daily enjoyment of the wireless depends upon this fact in that the Earth's ionosphere prevents the radio waves which we want to receive from escaping into the space outside; also we are sometimes robbed of this enjoyment when solar activity produces ionization in inconvenient parts of the Earth's atmosphere.

In much the same way as the Earth's upper atmosphere is transparent to ordinary light but opaque to ordinary radio waves, so the Sun's outer atmosphere is transparent to light but opaque to short radio waves. But a body which is opaque to any particular radiation behaves, so far as that radiation is concerned, as a black body and itself emits that radiation with an intensity corresponding to its own temperature. Therefore the solar corona may be expected to emit radiation in certain radio wavelengths as though it were a black body at whatever temperature it possesses; we have seen that this temperature is estimated from entirely independent evidence to be of the order of a million degrees. Thus we reach a possible explanation of the radiation actually observed.

The calculations suggested by this discussion have been made by several workers, notably by D. F. Martyn. He finds good quantitative agreement between theory and observation as regards both the intensity of the radiation and its dependence upon wavelength. The explanation is in fact generally accepted. What it means is that the human eye, being sensitive only to ordinary light, sees through the corona as though it were not present and places the apparent surface of the Sun at the top of the photosphere. Were the eye sensitive to any other wavelength interval up to a centimetre or so, it would see the same surface. But were the eye sensitive only to wavelengths between a few centimetres and a few metres, it would not see through the corona but it would see the Sun as possessing a rather fuzzy "surface" somewhere in the corona itself.

It must finally be pointed out that, even though the solar noise is so very much greater than that corresponding to a solar temperature of 6,000 degrees, the total *energy* emission in radio wavelengths is exceedingly small compared with the energy emission in visible light. Interesting as these phenomena are in themselves, they are energetically much too feeble to have any influence upon the behaviour of the Sun as a whole.

Disturbed Sun. Sunspots and solar flares produce outbursts of radio noise reaching peak-intensities many times greater than that of the quiet Sun, sometimes as great for the wavelengths concerned as would correspond to a solar temperature of

about 10^{10} degrees. Actually it was the noise associated with the passage of spots across the solar disk which was first shown by J. S. Hey in 1942 to have definitely a solar origin, and the discovery of solar noise under other conditions followed later.

Hey and others have shown that the radio emission associated with a sunspot is apparently a "beam" emission radial to the Sun which produces the observed outburst of noise as it sweeps across the Earth. The emission from a flare, on the other hand, does not appear to be beamed, and lasts only during the brief time of activity of the flare.

Of the many attempts to explain these emissions none has yet met with general acceptance. It is not even definitely established that the emissions come directly from the spots and flares with which they are associated; it has been suggested that these disturbances in the photosphere may merely cause enhanced emission from the overlying parts of the corona.

Galactic noise. It is convenient to mention this phenomenon here, since not enough is known about its production to enable us to relate it to any of the later topics.

Observation shows that the galactic system produces radio emission, called *galactic noise*, in the same wavelengths as those of solar noise. Apparently some noise comes from most parts of the Galaxy, but several regions have been shown to give specially intense radiation. There have been found, moreover, what appear to be point sources of noise and these cannot be identified with any visible features of the galactic system. Lastly, the radiation from some of the regions of strong emission is found to include components of fluctuating intensity having periods of the order of minutes or seconds.[1]

Practically all that can be said with certainty is that, if the galactic noise is produced by stars, then some stars must be vastly more productive of noise than the Sun. Attempts have

[1] Since the above was written, tentative identifications of certain of the "point sources" with other objects in the Galaxy have been announced. Also, it has been fairly conclusively established that most, if not all, the fluctuation of intensity is due to the action of the Earth's ionosphere, and is not inherent in the sources themselves.

been made to account for the noise by free-free transitions in interstellar gas. This agency may account for some fraction of the observed radiation but is unlikely to explain everything.

Radio astronomy is a very new subject and may be expected to develop rapidly in the near future. Considerably more complete observational material is required before much further theoretical progress can be expected. In particular, there is a need for appliances, at present lacking, for obtaining spectra of radio noise. On the theoretical side, it is scarcely to be expected that convincing theories of galactic noise will emerge until the mechanism of production of solar noise associated with solar activity has been discovered.

Solar Observations

This book is not intended to deal with observational technique, but it ought to be stated here that the observational knowledge of the Sun which we have been quoting has been secured by the use of many appliances contrived specially for solar work. This is particularly true of the observation of all forms of solar activity. It applies most of all, however, to observations of the corona which until about 1930 had been observed only during the brief duration of solar eclipses, but can now be observed at all times (under favourable conditions) by means of the ingenious devices first developed by B. Lyot. Actually, such devices are likely in the long run to prove most important for facilitating the detailed study of the whole solar atmosphere and are being developed to that end by Lyot and other astronomers.

It ought also to be mentioned that, in addition to the study of solar radiation, in the form of radio noise, in wavelengths far removed from the visible range, our knowledge of the radiation is now being extended into ranges immediately adjoining the previously accessible range. American astronomers have obtained photographs of the spectrum in the far ultra-violet, using apparatus carried by rockets into the Earth's upper atmosphere. In the infra-red, the spectrum has been studied by using lead sulphide or selenide cells, in place of photographic plates, to record the radiation.

CHAPTER VII

THE SUN: THE SOLAR INTERIOR

So far as the study of the Sun as a whole is concerned, all the work so far described is the merest scratching at its surface. It concerns only a few millionths of the Sun's material. We turn now to an attempt to study the main body of the Sun. We want to discover the physical state of the matter inside the Sun and to find out what mechanism it is which keeps the Sun shining.

Fundamental data. Let us first set down what we can claim to know as the foundation of our investigation.

A. *The Sun is a spherically symmetrical body of known mass M and radius R held together by its own gravitation.* There are no external forces acting upon the Sun which can affect its internal structure: its own rotation is not fast enough appreciably to affect its spherical symmetry.

B. *The Sun is generating energy at a known rate L.* "Liberating" would be a better word than "generating", though the latter is more usual and is correct in the same sense as when we speak of a power-station generating energy when it is, more accurately, only liberating energy stored in water behind a dam or stored in some form of fuel. Accepting the conservation of energy, the Sun must be liberating energy from some internal store at the same rate at which it is radiating energy at the surface. This rate has been quoted in Table I; it is an average of 12 horse-power for every ton of material in the Sun.

C. *The Sun, to a high degree of approximation, is in a steady state both mechanically and thermodynamically.* This is a direct result of observation so far as it concerns the time during which the Sun has been accurately observed.

We have already pointed out that the occurrence of solar activity is not at variance with this assertion concerning the Sun as a whole. Also it has been inferred from geological evidence that the Sun cannot have changed much, if at all, in a thousand million years—and a body might change a lot in that time and still be regarded by all ordinary standards as being effectively in a steady state at any particular epoch.

Though it follows from item B that there must be some progressive change somewhere in the Sun, in that it is continually losing energy which it is not apparently replacing, we now conclude that the loss in any "ordinary" time-interval cannot have any appreciable effect on the state of the Sun.

D. *The boundary layers of the Sun have the properties described in the preceding chapter.* Though these properties concern such a small fraction of the total mass, they do, of course, provide the conditions to which the main mass of the Sun must conform at the boundary.

Internal conditions: preliminary estimates. We cannot know completely what the physical conditions are inside the Sun without knowing how it generates its energy. On the other hand we cannot expect to know how it generates its energy without having some general conception of the physical conditions in which the generation has got to operate. Fortunately a few simple arguments based upon the general statements of the last section will provide the necessary conception.

We start from the fact that what keeps the Sun distended against its own gravitation is its internal pressure. Let us imagine any plane drawn through the centre of the Sun and so dividing it in half. The two hemispheres are pulled together by their mutual gravitation and are kept apart by the pressure acting across the plane. Hence the total pressure is equal in magnitude to the gravitational pull. We cannot calculate this pull exactly without knowing how the matter is distributed. But we certainly expect the matter to be concentrated towards the centre, and the more it is concentrated the stronger is the pull between two halves. Now if the matter were not concen-

trated at all, i.e. if its density is uniform, a simple calculation shows that the pull would be about 10^{37} dynes (the *dyne* being the c.g.s. unit of force). The actual pull must exceed this amount. Therefore the average pressure over the diametral section, whose area is about $1{\cdot}5 \times 10^{22}$ sq. cm., must exceed $10^{37} \div (1{\cdot}5 \times 10^{22}) = 6{\cdot}7 \times 10^{14}$ dynes per square centimetre, or about 700 million atmospheres.

It has been said that the Sun's atmosphere consists largely of hydrogen. As a working hypothesis, we shall take this to hold good also for the interior. Now we know that the mean density of solar matter is 1·41 g. per c.c. or nearly one-and-a-half times that of water. If hydrogen of this density were to behave like a gas, then the elementary gas-law requires that, for a pressure equal to the average calculated above, the temperature must be about 3 million degrees. Under these conditions the hydrogen would be practically completely ionized and the value given for the temperature takes account of this.

At first sight, however, it appears absurd to contemplate matter denser than water behaving like a gas. For if normal hydrogen atoms were packed together to give the required density they would be in actual contact with each other, and so would not possess the freedom of movement which characterizes a gas. But this is *not* the state of affairs we have to consider. For, as we have said, the atoms would be ionized; so instead of normal atoms we have to consider free protons (hydrogen nuclei) and free electrons. So far as any size can be ascribed to these particles, the electron has a diameter about one hundred-thousandth that of the normal hydrogen atom and the proton is even smaller. Therefore, even at densities far greater than that under consideration, these particles have ample room to run about freely like the atoms or molecules of an ordinary gas. We conclude that the hydrogen does behave like a gas and that its temperature may in fact be calculated accordingly.

Moreover, the same general conclusion holds good whether we consider hydrogen or any other element or mixture of elements. Had we considered, for instance, iron instead of hydrogen we should have found a temperature of about 12

million degrees to be required to give the same pressure for the same density. Under these conditions the iron atoms would retain on the average only about three of their 26 electrons. These stripped atoms are vastly smaller than any normal atoms and once again the assumption of gas-like behaviour would be justified. Further, since a single normal atom of iron of mass 56 times that of a hydrogen atom would be replaced by some 24 particles (ionized atom plus about 23 free electrons) the average mass of these particles, or the *mean molecular weight* of the gas, is about 2·3 times that of a hydrogen atom. Approximately the same values of the temperature and mean molecular weight would apply for all but the lightest atoms. Thus the gas-like behaviour would follow whatever the chemical composition of the material, and the temperature and mean molecular weight depend only to a minor extent upon this composition.

We must recall that, as we saw in Chapter IV, whatever the temperature at any locality in the gas, the corresponding temperature radiation must be present. Now such radiation itself exerts a pressure: a standard formula shows that for a million degrees this pressure is 2·5 thousand atmospheres and for 10 million degrees it is 25 million atmospheres. Under the conditions we are discussing the radiation pressure is, therefore, small compared with the gas pressure and we are justified in neglecting it in our rough estimates. It is, however, on the verge of becoming relatively significant, and the possibility of its doing so must not be ignored in dealing with stellar interiors in general.

It should be pointed out that we have been using an average density and an average pressure for the Sun with rather different meanings for "average". If we now call the corresponding temperature an average temperature, it will be rather a hybrid kind of average. However, at this stage we are concerned only to discover the orders of magnitude and any significant average may be expected to give the same order of magnitude. On this understanding, we are now able to assert with considerable confidence:

The whole of the Sun's material is in gaseous form; *its average pressure exceeds about* 700 *million atmospheres*; *its average*

temperature exceeds some 3 million degrees; practically the whole internal pressure is gas-pressure, radiation pressure being relatively unimportant.

It is rather remarkable to be able to infer so much from no solar data other than the mass and radius; for the rest, the conclusions depend only upon general atomic physics. About twenty-five years ago, when Eddington began to announce such conclusions for the stars in general, they were regarded as revolutionary. He brought forward a formidable array of astronomical evidence to support them. But this was because he had to vindicate not only the astronomical conclusions but also to a great extent his application of the then novel results of atomic physics. It is not always realized how, now that the atomic physics has become more commonplace, a good general idea of conditions inside the Sun can be derived by quite simple arguments.

Variation from boundary to centre. At any distance r from the centre of the Sun the pressure P has to support the weight of material between r and the surface. Hence P increases as r decreases. Also the temperature T at distance r must increase as r decreases; were T to decrease inwards at any level there would be an inward flux of energy at that level so that the part of the Sun inside that level would be continually imbibing energy from the part outside, which would be impossible in a steady state. The inward increase of temperature must partly account for the increase of pressure. But we expect on general grounds that the density ρ must also increase inwards.

The measures of the variation of P, T, ρ with r will be called the *gradients* of these quantities. The temperature gradient at distance r depends upon the energy-flux at that level and the means of energy transport. The energy-flux is determined by the rate of energy-generation inside the level. So we cannot proceed much further without knowledge of the means of energy-transport and the mechanism of energy-generation.

There is, however, a line of reasoning similar to what was applied in the preceding section to average values which show that, were the central temperature of the Sun as large as 100

million degrees the central density would have to be about a thousand times the mean density. This would represent a fantastic concentration of mass towards the centre. So we are fairly confident that the central temperature must be well below 100 million degrees. But we know that it must be well above the average temperature which itself has been shown to exceed 3 million degrees. Consequently we may make a provisional estimate that the central temperature is *of the order of* 10 million degrees.

This rough estimate is important for the following reason: It gives, indeed, a fall of temperature of the order of 10 million degrees from the centre of the Sun to its boundary; but this fall takes place in about 700,000 kilometres and so it averages less than 20 degrees per kilometre. Further, 20 degrees is less than one part in 100,000 of the average temperature. Thus the proportional temperature gradient is minute: if it occurred in a laboratory it would be far too small to be even detected. Hence, to an exceedingly high accuracy, the material of the Sun is in local thermodynamic equilibrium (page 48).

Solar Energy-generation

Problem. The Sun has been shining at about its present brightness for some hundred million years and probably for a much longer time. Our problem is, How has the Sun maintained this output of energy?

Ever since man has understood that his material being depends upon the light and heat he receives from the Sun he has implicitly realized the importance of this problem. For some centuries, at least, he has explicitly recognized the existence of the problem and speculated upon its solution. But not until he had come to understand it in terms of energy and to accept the principle of the conservation of energy could he properly formulate the problem. Even then he could not consider it quantitatively until he had some estimate of the length of time during which the energy output has endured. Finally, he could scarcely hope for a solution until he knew the general nature of the physical conditions obtaining throughout the Sun.

The requirements for a serious attack upon what must be

ago. Physicists and astrophysicists believe that the problem has now been solved.

Possible energy-sources. Our normal domestic methods of generating heat are to burn something or to let something fall. The burning may be done on the household hearth or at a power-station; the falling is usually that which produces water power.

These two methods illustrate in fact the only two ways of obtaining energy from matter. Burning is a chemical reaction involving a *rearrangement of fundamental particles* into a system of less energy than the original ones, the energy difference between the original and final systems providing the heat of reaction. The energy is associated with the operation of *electrical* forces between the fundamental particles. Falling is a *rearrangement of matter in bulk* into a system of less energy than the original one. The energy is associated with the operation of *gravitational* force. Since there are no other ways of rearranging matter, and since fundamentally there are no other forces than electrical and gravitational (the theory of relativity having abolished any fundamental distinction between magnetic and electric force) there are no methods of generating heat essentially different from the two familiar ones mentioned.

Unless the laws of physics as we know them cease to apply inside the Sun we have therefore to seek its source of energy-production only in some form of "falling" or some form of "burning".

Let us take "falling" first. The only kind of falling the Sun can do is to fall in upon itself, i.e. simply to contract. The hypothesis that the Sun by gradually contracting is all the time converting gravitational energy into heat energy was the famous one advanced last century by H. v. Helmholtz and Lord Kelvin. At the time it seemed to offer the only possible solution. But, even were the contraction taking place, it would

not answer for more than about one per cent of the total energy output required. Eddington calculated "that 20 million years is probably a generous estimate of the Sun's age on the contraction hypothesis".

There had been an earlier hypothesis which made the Sun shine by the release of gravitational energy: it ascribed the Sun's heat to a bombardment by meteors falling on to the surface. But, if the falling in of the whole mass of the Sun cannot supply enough energy, then the fall of meteors, which would necessarily add up to less than the whole existing mass, certainly cannot do so.

It is therefore established that gravitational energy is not the source of the Sun's radiation.

We conclude therefore that the source is some form of "burning". But burning in the usual sense of the term necessitates the existence of chemical *molecules*, possibly as the combustible and certainly as the combustion-product. Our inferences about physical conditions inside the Sun show, however, that no molecules could possibly exist there. As it has been expressed, the material of the Sun is too hot to burn! In any case, even were the Sun a mass of pure carbon, its complete combustion would supply enough energy to maintain the Sun's radiation for only a few thousand years. The same would be true for any other combustion process. Thus the amount of energy released by such processes would be altogether too paltry even were it in any way possible for them to occur in the Sun.

By this process of elimination we are left with only one further possibility. If the material of the Sun cannot burn so as to produce new chemical compounds, can it burn so as to produce new chemical elements? In other words, if the fundamental particles cannot rearrange themselves into molecules, *can they rearrange themselves into new atoms*? Logically there is no other possibility and, even if laboratory physics had not revealed the transmutability of the elements, solar physics would have demanded it.

Subatomic burning. In ordinary burning we have to do with reactions that result in the making or breaking of chemical

the atoms remain unimpaired.

It is well-known, however, that reactions between nuclei in fact possible and result in the *transmutation of elements*. re the relatively insignificant affairs of the electronic systems y be ignored. We have to do, not with federations, but with nplete mergers. Two nuclei lose their identities in the forma- n of a new one; in some cases the latter may be an unstable ucture which soon sunders itself into more stable com- nents. In any event, the final result is one or more new clei in which the individualities of the original ones have ite vanished.

The energies of such subatomic reactions are of an alto- her greater order of magnitude than those of the most ergetic ordinary chemical reaction. The comparison is rally that of an atomic bomb with an ordinary high explosive mb.

It happens that, just because the energies are so great, y can be predicted for any specified reaction in a simple d accurate manner. The principle which renders this ssible is that commonly called the "equivalence of mass d energy". This was originally a theoretical inference in the eory of relativity, but it is now claimed as a well-established perimental result. It asserts that, if any system gains or es an amount of energy E it thereby gains or loses an amount mass E/c^2, where c is the velocity of light, both quantities ing measured in the usual units; conversely, if a system ins or loses an amount of mass M, otherwise than by the nsfer of matter, it thereby gains or loses an amount of ergy Mc^2.

Remembering that c is about 3×10^{10} cm./sec., so that e mass associated with one erg is about 10^{-21} gramme, it is ident that the mass-changes associated with everyday ergy-changes are utterly insignificant. This is not the case r nuclear transformations. Consider, for instance, a helium

nucleus or α-particle as we can call it. Its mass is so nearly equal to that of four protons that we are compelled to examine the hypothesis that it is somehow formed by the union of four protons. But the accurately determined masses show that the mass of the α-particle is in fact not four but 3·97 (approx.) proton masses. Consequently, if four protons do combine to form an α-particle there is a loss of $4-3{\cdot}97 = 0{\cdot}03$ (approx.) proton mass. We conclude that an equivalent amount of energy must be liberated. This energy would then be the "heat of reaction" of the process, though we should expect it first to reappear in the form of radiation.

The best way for us to appreciate the implication of this conclusion is the following: Suppose the present mass of the Sun to be due entirely to protons and then suppose the protons to be completely transformed into α-particles. For every four protons there would be a loss of about 0·03 proton mass. So a fraction about $0{\cdot}03 \div 4 = 0{\cdot}008$ of the whole mass of the Sun would be converted into its energy-equivalent. This quantity of energy would suffice to maintain the Sun's present rate of radiation for about 130,000 million years.

Looking at it another way, if the Sun has been radiating at its present rate for the past 2,000 million years, and if it was originally composed mainly of hydrogen, then its total energy-output during the whole of that time could have been achieved by the conversion into helium of about 1½ per cent of its hydrogen.

We have thus found a type of process which would tap an energy-supply wholly adequate for the Sun's estimated requirements. We shall shortly come to the vital questions: Do such processes actually occur? and, if so, Do they proceed in the Sun at the rate required to maintain the observed rate of energy-output?

Other subatomic processes. It is convenient first of all to explain why our problem cannot be solved in terms of the more generally familiar subatomic processes of natural *radio-activity* and *nuclear fission*. The latter is the process employed in the atomic bomb and in the proposed industrial uses of atomic energy.

As possible explanations of solar energy-generation these processes are ruled out by the extreme cosmic rarity of the raw materials, the radioactive and the fissile elements. But even if it were supposed, contrary to all the evidence, that the elements are sufficiently abundant to provide the required energy-generation in the Sun and stars, then it could easily be shown that this would lead to inconsistencies with other known facts about the behaviour of these bodies. So the rarity of the elements would only receive added confirmation.

Thus, although current discussions about atomic power have popularized a considerable amount of knowledge of atomic nuclei and their behaviour, such knowledge mostly concerns types of behaviour which appear not to be of astrophysical significance. The only common features in the processes involved in existing projects for the terrestrial uses of atomic energy and in the solar processes we are about to describe are the very general ones that nuclear particles do undergo transmutation and that they do thereby release great amounts of energy. This is not to say that man will not at some future time succeed in reproducing on an industrial scale some of the nuclear processes which occur in the Sun.[1]

Thermonuclear processes in the Sun. The formation of helium out of hydrogen, which we cited merely in order to illustrate the quantities of energy available from nuclear reactions, is in fact not a process which can occur in a single reaction. Apart from the circumstance that the formation of an α-particle (charge-number 2) out of four protons (each of charge-number 1) would require the disappearance of two units of electric charge, it would require the simultaneous interaction of four protons. But it can be shown that, even at the greatest densities expected to occur in the Sun, the chance of four nuclei all interacting absolutely simultaneously is entirely negligible. The only type of interaction which it is worth while considering is that of only *two particles at a time.* (Compare the likelihood of an accidental collision between two aircraft

[1] The processes contemplated in connexion with the recently discussed possibility of a "hydrogen bomb" are probably more akin to the solar processes than are the processes of radioactivity and nuclear fission.

with that of four aircraft all colliding accidentally at precisely the same instant.)

Now it has been shown experimentally that, when a proton collides with a heavier nucleus it may under certain circumstances penetrate the latter. Then this either forms a single new nucleus or else splits into two new nuclei each more complex than a single proton. The measured masses of the particles involved shows that such a process causes the disappearance of about 0·008 of the mass of the proton. This is about the same as would be the loss per proton were the formation of an α-particle from four protons a possible process. An amount of energy equivalent to this mass-loss must be released. It will appear first as the energy of a photon, or as kinetic energy of the resulting particle or particles, or as shared between these forms. But, if the process occurs in the Sun, in whatever form the energy first appears, it will quickly be converted into ordinary thermal energy of the matter and radiation.

Quantum mechanics shows how to calculate the probability that a proton H of given energy striking a nucleus[1] Q will be captured by Q. Empirical data exist which fix the constants in the formula. The latter shows the probability to decrease rapidly with increasing charge number of Q in such a way that we need to consider only cases where this is less than about 8.

When we speak of a proton "striking" the nucleus Q we imply that Q presents a definite target area in respect to this particular form of attack. This area can be estimated from empirical or semi-empirical data on nuclear processes. Knowing the target area, it is an elementary calculation in the kinetic theory of gases to determine the number of relevant collisions taking place per unit volume and unit time in stellar material of any assumed composition, density and temperature. The expression for this number can be made to show its dependence upon the relative kinetic energy of the participating protons.

Combining the formulæ for the collisions and for the probability of capture in a collision, we then determine the

[1] We use Q to denote some specified nucleus; it is not the chemical symbol of any particular element.

rate at which the capture process is proceeding in the given material. It is in fact clear that the rate is proportional to the numbers of H and Q nuclei present per unit volume and for the rest depends upon the *temperature* T: the higher the temperature the faster the particles move in their thermal agitation and consequently the more frequent and energetic are the collisions. Hence the process is called a *thermonuclear process*, thus emphasizing the distinction between it and the "spontaneous" processes of radioactivity which are unaffected by thermal conditions. The formula for the rate of reaction actually shows it to be extraordinarily sensitive to the value of T.

In the simplest case, the reaction product would be a single new nucleus R, say, of mass and charge each one unit in excess of those of Q. However, R may or may not exist as a stable nucleus; if it does not, then it spontaneously decays in some particular manner. We may cite some examples:

If Q is itself a proton, so that we have the simplest possible nuclear reaction—the proton-proton reaction, then the end-product is not a nucleus of mass and charge each two units. It is a deuteron (nucleus of "heavy" hydrogen) and a positron (positive electron).

If Q is a deuteron, and if it can capture a proton, the immediate result would be a nucleus of the helium isotope of mass 3. In contrast with the first case, this is believed to be a stable structure and so would be regarded as the end-product.

If Q is a lithium nucleus of mass 7, the immediate result R is a nucleus of the beryllium isotope of mass 8. This is known to be unstable and to disintegrate into two α-particles which therefore constitute the effective end-product of the reaction.

In all cases where R is not stable its subsequent disintegration must release some additional energy. However, this is in general small compared with that released in the formation of R itself.

It is evident that a formidable collection of laboratory data on the properties of atomic nuclei is demanded in order to obtain quantitative results concerning thermonuclear reactions in stellar material. Though the possible significance of such reactions had been appreciated for some years, notably on

account of pioneering suggestions advanced by R. d'E. Atkinson in 1931, sufficiently definitive data were not forthcoming until about 1937–9.

In 1938, C. F. von Weizsäcker and H. A. Bethe independently surveyed the available facts and arrived at similar conclusions regarding the problem before us. The nuclear phenomena which have been most intensively studied in the intervening years have not been immediately relevant to this problem. Doubtless much fundamental information has accumulated which may ultimately find astrophysical application. At the moment, however, knowledge of nuclear processes in the Sun is very much what it was in 1939. The main conclusions of von Weizsäcker and Bethe concerning these processes have in any case been generally accepted and we shall now consider them.

We remark first that only processes involving proton-capture need be considered, since in any others the amount of mass which would disappear would be too small to make a significant contribution to the Sun's energy-output. Therefore what is needed, and what Bethe carried out, is an exhaustive survey of the potentialities of H-Q reactions, as basic reactions for solar energy-production, for all possible nuclei Q. If Q is any nucleus having a mass-number[1] from 2 to 11 inclusive Bethe concluded that the H-Q reaction is unimportant for the process. It could give a great deal of energy, but it would proceed too quickly and all the available Q nuclei would have long ago been burned up in the reaction. Therefore, if these light nuclei exist at all in the deep interior of the Sun it can be only because they are being continually reformed as intermediate products in *chains* of nuclear reactions.

There are left for consideration the H-Q process in which:

A. The mass-number of Q is 1, that is the proton-proton reaction;
B. The mass-number of Q is 12 or more.

In process "A" the only raw material needed is protons.

[1] The mass number is the nuclear mass expressed as the nearest whole-number multiple of the proton-mass.

ıs into an α-particle (helium of mass-number 4), if the rature is high enough, or the combination of three ıs into a nucleus of the helium isotope of mass-number 3 er temperatures. This is because every possible chain to some unstable nucleus such as beryllium 8 which grates into such products (except for insignificant ers of other surviving products).

ıe very important by-product of this discussion is the sion that *progressive* atom-building does not take place ıormal star. Were the star to consist initially of pure gen, this hydrogen would gradually get transformed into ı. At any instant the material would consist almost y of these two elements; there would be a small propor- other light nuclei but these would only form the material is in course of transition from hydrogen to helium. might be a few heavier nuclei which by rare chances survived intermediate stages of instability, but their er would be negligible. Thus a *normal star*, if it consisted ally of pure hydrogen, *cannot generate within itself a e of chemical elements* such as we find on the Earth or solar atmosphere. Since a normal star undoubtedly ns such a mixture, it must either have been endowed t from birth, or have acquired it from outside, or have ated it during some abnormal phase of its career.

case "B" the following chain of reactions is possible, ımbers being the mass-numbers of the isotopes concerned:

carbon 12	+ proton	→	nitrogen 13	+ photon
nitrogen 13		→	carbon 13	+ positron
carbon 13	+ proton	→	nitrogen 14	+ photon
nitrogen 14	+ proton	→	oxygen 15	+ photon
oxygen 15		→	nitrogen 15	+ positron
nitrogen 15	+ proton	→	carbon 12	+ α-particle

This consists of four H-Q reactions and the two spontaneous transitions of nitrogen 13 and oxygen 15. The Q nucleus in the first reaction is carbon 12 and the end-product of the last reaction is a nucleus of carbon 12 and an α-particle. Thus the net result of the chain is simply that four protons have disappeared and one α-particle has appeared with the release of some radiation (the photons) and two positive electrons (positrons). By virtue of the conservation of energy, the total energy released in the chain is the same as if the protons had combined directly to produce the same result.

The highly significant feature is that the nucleus of carbon 12, which disappears from the scene in the first reaction, is replaced by an identical nucleus in the last one, while each of the other heavier nuclei which appears in any one of the reactions promptly disappears in the next. Thus, unlike those reactions which start with a nucleus Q of mass 2 to 11, the Q nuclei in the present sequence are not burned away; after entering into a number of temporary combinations, they are finally reconstituted in their initial state. In chemical language the carbon operates merely as a *catalyst* for the transformation of hydrogen into helium.

As regards all the nuclei, other than protons and α-particles, the chain of reactions is purely *cyclic*; it may be regarded as starting anywhere in the above sequence. It is the now well-known *carbon-nitrogen* (C-N) cycle. Bethe has shown that it is the only such cycle which need be considered for our purpose; the charge-number of oxygen is 8 and we have mentioned that H-Q reactions in which Q has any charge-number greater than this, and which would have to be included in any other cycle, are too improbable for our purpose.

The importance of the C-N cycle is that, if a star consists initially of almost pure hydrogen but contains a small proportion of any of the heavier nuclei involved in the cycle, then the cycle can and must proceed, these nuclei being used over and over again in the catalytic conversion of the hydrogen into helium with the continual release of energy resulting from this conversion.

In such a star we have therefore two simultaneous processes which both achieve as the same net result this conversion

of hydrogen into helium: Process "A" which starts with the proton-proton reaction and which we regard as the direct synthesis of helium, and Process "B" which is the catalytic synthesis with the aid of the C-N cycle.

Rate of energy-generation. Bethe calculated, in accordance with the principles we have explained, the rates of energy-generation in the Sun resulting from processes "A" and "B". He adopted the values of the temperature and density previously estimated by Eddington (with which the rough estimates given above are in agreement so far as they go) and he took the chemical composition of the material to be that previously estimated for the Sun's atmosphere by H. N. Russell. His resulting estimates for the rate of energy-generation per gramme per second averaged through the whole Sun were: Process "A", about 0·2 ergs; Process "B", about 3 ergs; actual known rate for the Sun, 1·9 ergs.

The immediate conclusion from these figures is—that Bethe was incredibly lucky in the solar conditions he assumed! For he also calculated that, in the temperature region considered, the rate of energy-generation by process "A" varies approximately as $T^{3.5}$ and by "B" approximately as T^{18}. So had he assumed a temperature, say, twice as great as he did he would have calculated a rate of energy-generation about 2^{18}, or more than quarter of a million, times the value he gave! And in this case process "B" would have given a rate about 20,000 times greater than "A".

What Bethe had actually discovered were two processes, both of which can and must proceed in the Sun if his arguments are valid, and each of which could give the required rate of energy-production *if* the temperature and density have certain values close to those actually believed to apply. But, if these values were chosen so that process "A" gives the required rate, then they would be such that process "B" would give far too vast a rate. Hence process "A" cannot be the important one. Further, since Bethe's arguments have been found so convincing, we can now assert that the temperature and density *must* be such that the C–N cycle gives exactly the observed rate of generation.

Remarks on C–N cycle. As remarked towards the end of Chapter IV the condition at any locality inside the Sun where the C–N cycle is in operation may be regarded as that of quasi-thermodynamic equilibrium. The energy-generation has to be looked upon as a "slow reaction proceeding under conditions of (local) thermodynamic equilibrium". How slow the reaction is can be seen from the fact that at the centre of the Sun each cycle takes a few million years to complete; the high rate of energy-generation is due to the large number of cycles going on at the same time.

Recalling the principle of exact reversal stated in Chapter III, we realize that the present nuclear processes could all be reversed. But it can be shown that the reverse processes would not operate at rates comparable to those of the direct processes at temperatures below about 1,000 million degrees: so they can be neglected in the Sun.

Finally, there is the fate of the two positrons created by each cycle to be considered. All we need say is that in the Sun a positron quickly makes an encounter with a free negative electron in which both are annihilated. The energy thereby released can be regarded as reckoned in the total output of the cycle. The annihilation of free electrons by this means keeps the total number of electrons in correspondence with the total nuclear charge remaining.

Internal Constitution of the Sun

Problem. We have seen how the physical conditions inside the Sun can be roughly but confidently inferred merely from its known mass and size. We have also seen that only sub-atomic reactions can be expected to supply the Sun's energy-output. Following Bethe's detailed examination, we have seen that one and only one particular such reaction is almost certain to supply the known output under the inferred conditions.

We have, so to say, taken the Sun to pieces to see how it works and we are fairly sure we have got the answer. But the real test is, Can we put the Sun together again and make it work? To be a little more technical, The Sun has known mass M; *if* we use its known radius R, we can fairly safely say what

of the Sun.

The problem before us can be put this way: Suppose there is isolated in space a quantity of matter of mass M. It will pull itself together by its own gravitation and form something that can be called, quite non-committally, a "star". As it pulls itself together, or falls in on itself, some of its gravitational energy will be converted into heat energy and some of this will be radiated away. If nothing happens to stop it, this gravitational contraction will go on. But the tendency of the contraction is to produce higher and higher internal temperatures in the mass. So, sooner or later, sub-atomic reactions must come into operation and must start releasing sub-atomic energy. There must come a stage when the rate of release is exactly equal to the rate of radiation from the surface. We must expect this stage to be approached in such a way that, when it is reached, the "star" will steady up into what can be regarded as an equilibrium state. The problem is, What are the characteristics of this state; in particular, what radius has the "star" then attained, and what rate of energy-generation, i.e. what luminosity?

The physical theories we have endeavoured to describe in this book, when put in their technical forms, supply all that is needed for the solution of this problem. For the rest it is a mathematical exercise. The mathematician has, in fact, to construct a (theoretical) working model of the Sun. The physicist tells him merely the properties of the material he is to use in his construction. The astronomer tells him merely the total quantity of material he may use. When the mathematician has produced his model, the astronomer compares it with the real Sun.

The situation is rather analagous to what might happen in war-time when the enemy starts using a new weapon, say a rocket. One can imagine service representatives calling in a mathematician and saying, "We estimate the overall weight

of the rocket to be so much; we believe the enemy to be usin such and such materials to make it; now you work out fc us what sort of a rocket of that weight can be made of thos materials and what its performance must be". The mathe matician looks up the properties of the materials which hav been mentioned and proceeds to construct, on paper, h theoretical rocket. The service people then compare i theoretical performance with what they know of the actu performance of the enemy's weapon. If there is satisfacto agreement, then they may claim to know how the actual rock is constructed and can predict things about its performan which have not yet been learned from actual experience.

The mathematician in this hypothetical case will, howeve almost certainly produce a whole series of theoretical rocke corresponding to different ratios of weight of propellant weight of explosive, what he may call briefly rockets of vario "composition". One of the things the service people shou then be able to do by comparing his predictions wi performance is to pick out the particular theoretical rock which best fits the facts and so to discover what actu "composition" the enemy is using.

The last feature has an exact counterpart in the sol problem. The astronomer can tell the mathematician the to amount of material in the Sun, but not its chemical compo tion, i.e. the proportions of the different chemical elemen So the mathematician has to supply a series of models f various assumed compositions. If one of these gives t correct size and brightness of the Sun, then it may be tak to give also the correct composition, and this will be one of t most important results of the whole investigation.

The mathematician can foresee by inspection of the formu for the physical properties of the material that his results w be sensitive only to the proportions of hydrogen, of helium, a of the rest of the elements all taken together. What eleme the "rest" consists of is scarcely significant. The reason for t somewhat surprising situation will become plain in Chapter I

Results. The mathematical problem has been solved by large number of workers. As a result of their calculatio

To be as cautious as we need, we can say at any rate that these calculations and those for stars similar to the Sun leave no reasonable doubt that the energy is generated by a sub-atomic reaction whose dependence upon density and temperature must be of the same general character as for the C–N cycle. All the nuclear evidence available is in favour of its being actually the C–N cycle.

The calculations show, further, that the interior of the Sun consists of two zones: (a) the central core which is in *convective* equilibrium, which contains about 12 per cent of the total mass, and within which effectively all the energy-generation takes place. The central temperature is about 20 million degrees and the central density is between 50 and 100 times the mean density of the whole Sun; (b) the remainder of the interior forming a region in *radiative* equilibrium. The two regions merge into one another, but the transition takes place in a relatively thin layer.

These features are common to all current solutions of the problem. The current estimates of the chemical composition vary considerably. One set of estimates puts the hydrogen content at about 80 per cent, by numbers of atoms, the helium at about 20 per cent, and the heavier elements at about 1 per cent. Another puts the hydrogen at nearly 100 per cent, the helium at about 1 per cent, and the heavier elements at something very much less than 1 per cent. On the first view the opacity of the radioactive zone is *photo-electric* (i.e. due to bound-free transitions); on the second, the photo-electric opacity is appreciably supplemented by *electron scattering* opacity. In this respect, more work is required before a definite conclusion is reached. But there is no doubt that hydrogen is much the most abundant element throughout the Sun, as it has already been seen to be in the Sun's atmosphere.

CHAPTER VIII

STARS: DATA AND CLASSIFICATION

A QUITE rudimentary study of the general facts known about the stars is sufficient to show that they are objects of the same general character as the Sun. That is to say, each is an effectively independent body held together by its own gravitation and emitting radiation from its own resources. Also the characteristic quantities associated with the stars are such that the values of these quantities for the Sun are well within the ranges of values for the stars in general. It turns out, indeed, that the Sun is a "pretty average" star in almost every respect. Consequently, it is not surprising to find that the observational data available for the study of the stars are of the same character as those we have for the Sun, though for any individual star[1] they are less detailed and precise. It seems true to say that no observations of what may be called the *static* properties of the stars raise any fundamentally new problems not met with in the case of the Sun.

Where new problems do arise is in connection with the variability of certain stars (using this term to include the phenomena of novæ). Stellar variability, when properly understood, will certainly help also in the better understanding of stars which are not classed as variables and which are in a majority. However, in the present state of knowledge, it is convenient to exclude variable stars from the discussions in this and the succeeding chapter.

The data we are about to survey are mostly provided by stars in the nearer regions of the galactic system. But there is every reason to believe that, so far as our purposes go, these stars form an adequate sample for the whole of the Galaxy. We are interested here only in the stars as individual bodies and the data concern their masses, sizes, luminosities, spectral energy-curves and line-spectra. We take these in turn, giving

[1] In this chapter we mean by "star" a star other than the Sun.

in each case some general remarks upon the principles of the determinations, then a summary of the results, followed by brief comments upon them. This chapter summarizes observational results; the next will discuss the physics of the stars.

Stellar masses. We know the mass of the Sun only because we can measure its gravitational pull on the Earth or some other member of the solar system. Similarly, we can know the mass of a star only from its gravitational pull upon some other body. Of course, even if any other star does possess a planetary system we cannot hope to observe the planetary bodies directly and so to use their motions to infer the mass of the star. Fortunately, however, there are a great many cases where a star possesses a companion body in the form of another star and the two revolve around each other under their mutual gravitational attraction. Indeed, probably more than half the stars are members of such combinations of two or, occasionally, more stars. If the circumstances of the orbital motion can be sufficiently fully observed, then the mass of each star in a binary system can be inferred. Somewhat less complete knowledge of the motion may provide information concerning the combined masses or the relative masses even though it does not give both. The essential point is that we know the masses of certain stars almost as directly as we know that of the Sun, and have useful information about the masses of certain other stars.

A word should be said here about a common feature of observational astronomy. As soon as astronomers have "direct" measurements of a particular quantity "A" in a sufficient number of cases they look for some empirical correlation between its value and some other observable characteristic "B". If, then, the characteristic "B" can be observed in further cases in which "A" cannot be measured directly, the empirical relation may be employed to infer "indirect" values of "A" in these additional cases. Stellar mass is an instance of such a quantity "A". Without going into details, it may be stated that there exist more or less reliable estimates of the masses of many more stars than those for which direct determinations are possible.

Nevertheless, such indirect estimates are for the most part useful only for statistical studies or for general verification of certain theories. As fundamental data for the study of the physical constitution of the stars we ought, so far as possible, to rely upon only the most direct available determinations of the various quantities required. Consequently, the consideration of indirect data is largely outside the scope of this book.

A couple of further points about the determination of stellar masses require brief mention. While the assertion that we can know the mass of a star only from its gravitational pull upon some other body remains valid, the other body may in principle be a photon. For Einstein's relativity theory predicts that a photon must suffer a certain calculable loss of energy in escaping from a gravitational field. This produces the gravitational red-shift of spectral lines which constituted one of the three famous "crucial tests" of Einstein's theory. The validity of the result is now generally accepted. It means, that, if the effect can be observed for any particular star, it provides a "direct" determination of the star's surface gravity. The mass can then be determined if the radius is known, or *vice versa*. However, the effect is unmistakably detectable only in the most extreme cases of a few white dwarfs (see below), where it does in fact give valuable confirmation of values got by other means, and of a few exceptionally massive stars.

The "other body" of the above general statement may also be the star's own atmosphere. For the star's surface gravity is one factor affecting the pressure-gradient in its atmosphere and the pressure-gradient influences the spectrum. So we should expect the surface gravity to have a calculable effect upon the spectrum. If the effect is observable it ought to yield information about surface gravity and consequently about the mass. All this is in principle borne out by experience, but the effect is not usually regarded as a means of *determining* the mass and is best treated in another context ("Absolute magnitude effect", below).

Results. The smallest known stellar masses are about one-fifth the mass of the Sun. A few stars are known to have masses over 100 times that of the Sun, the greatest known

mass being about 400 solar masses. But *the great majority of known stellar masses lie within about* 0·4 *to* 4 *times the solar mass, i.e. a range of about* 10 : 1.

Comments. Compared with the enormous ranges for other characteristics of the stars, this is a very small spread of values. It has long impressed astronomers as probably having some fundamental significance.

It is now generally believed that the mechanism of star-formation tends normally to produce objects of a standard mass, within about an order of magnitude. Therefore, while we hope that our discussion in the next chapter will tell us why bodies of stellar masses behave as they do, we do not expect it to tell us why they possess those masses. We shall learn, however, that bodies of mass less than about one per cent of the solar mass would not behave like stars. So the range of stellar masses could not in any case be indefinitely extended in the direction of small masses.

Stellar radii. Suppose the distance of a particular star is known. Its size could be measured directly if its apparent angular diameter could be measured. But the biggest apparent angular diameter of any star is probably much under one ten-thousandth of one degree; for the vast majority of stars it can be only a minute fraction of this amount. Now an apparent diameter of one ten-thousandth of a degree is far too small to be measured by any of the usual methods for measuring angles.

What has just been said explains why, for direct observation, a star always behaves as a point-source of light. Every star is much too remote to provide an observable "disk" such as that provided by the Sun, Moon and planets.

Nevertheless, an ingenious method originated by A. A. Michelson does exist for measuring angles down to a few millionths of a degree. Light entering a telescope in precisely one direction can be made to interfere with itself and to produce an interference pattern at the focus. Light entering simultaneously in a very slightly different direction will produce a similar pattern slightly displaced in regard to the first. Using a null method, this effect yields a measurement of the small

angle between the two directions. Applied to a star it will, in principle, yield a measurement of its apparent angular diameter. The method has been successfully applied in practice to a few stars. But the overwhelming majority of the stars must remain forever outside the capability of this, the only method which may be described as a "direct" measurement of their sizes.

Fortunately, double stars once again come to our aid and in certain special cases provide an extremely convenient method of inferring accurate values of stellar diameters—the most accurate which we possess. The special cases are those of "eclipsing binaries", i.e. double stars whose orbits happen to be seen edge-on, or nearly so.

Consider, for definiteness, a case where the orbit is precisely edge-on and one star is appreciably bigger than the other. Then, once in each revolution the smaller star is lost to sight behind the bigger. When this occurs, the luminosity of the pair is reduced from the sum of the luminosities of the two components to that of the larger component alone. The time required for this reduction to take place is the time taken for the smaller star to travel a distance equal to its own diameter, while the duration of its total eclipse is the time for it to travel a distance equal to the difference of the two diameters. These times are easily observed from the "light-curve". If, as is sometimes the case, the actual speed of revolution of the pair can also be measured from the Doppler effect in the spectra, then these times can be translated into distances travelled and thus the absolute dimensions of the system can be inferred. The method has so far been successfully applied to some 50 eclipsing binaries.

By far the most abundant data on stellar radii are derived, however, by the method of comparing the total luminosity of a star with its luminosity per unit area. This will be explained later (page 115). The results are less accurate than those derived from eclipsing binaries.

Results. It is found that the radii of the stars range from less than $\frac{1}{100}$ that of the Sun to about 500 times that of the Sun, i.e. from less than the Earth's own radius up to something of the order of the radius of Jupiter's *orbit*.

material has an average mass of 20 tons, while there are others whose mean density is about one-thousandth the density of ordinary air.

Main sequence, white dwarfs, giants. Inspection of the data shows that there is not a steady gradation of the stars through the million-million-fold density range just mentioned. This fact is exceedingly important since it furnishes, in the approach we are making, the first indication of the possibility of *classifying* the stars.

Among all the stars for which good data are available, it is found that the majority form a set occupying only a relatively small central interval in the full density range. In fact, most members of the set have densities between about 10 times and $\frac{1}{10}$ the solar value. This set of stars is called the *main sequence* and we shall see that the name has been aptly bestowed.

The stars possessing the extremely great densities mentioned form a set quite unmistakably separated from the main sequence. They are called the *white dwarfs*.

Finally, the stars of very low density are *giants*. The density criterion makes it natural to treat these as a set to be distinguished from the main sequence, but this criterion alone does not yield a clear demarcation between the two sets. The further characteristics we are about to consider will clarify and amplify the classification, but it is instructive to have noticed that such a simple characteristic as the density shows the existence of more than one genus of stars.

Stellar luminosities. On the basis that mass and size are physically more rudimentary, we have deferred till now the observationally more immediately obvious characteristic of a star, which is its *brightness*.

We suppose we are dealing with stars whose distances are known. So, if the apparent brightness of the star is measured,

Provided the brightnesses have been observed for light in a fixed frequency-range, the comparison of the absolute magnitudes of the stars gives the ratios of their true brightnesses *in that particular light*. For instance, if all the observations were visual, so that we should obtain what are called "visual" magnitudes, then the comparison would be for light within the frequency range in which the human eye is perceptive.

Results. Once again, there is an enormous range of values. The brightest known stars are of the order of 20,000 times as luminous as the Sun, while the Sun itself is of the order of 20,000 times as luminous as the faintest stars.

It is noteworthy that the main sequence itself extends over almost the whole range of observed luminosities, even though it occupies only a relatively small portion of the density-range.

The giants are found to be all very bright stars and the white dwarfs very faint. It should be noted that this is by no means inevitable, *a priori*, merely from the fact that the giants have large radii and the white dwarfs small radii.

Comments. For the astrophysicist the most fundamental characteristic of a star's radiation is the *total* energy-output in all frequencies taken together. This is expressed by what astronomers call the *absolute bolometric magnitude* (bolometric = radiation-measuring), using a conventional scale analogous to that for visual magnitudes.

Were all stars to possess the same spectral energy-curve then, if a given star is found to be, say, ten times as bright as the Sun in visible light, its total luminosity would also be ten times the total luminosity of the Sun. Thus the visual luminosities already mentioned would suffice for the comparison of total luminosities.

A star's spectral energy-curve is the complete specification of the *colour* of its radiation. Even as seen by the naked eye, the stars differ greatly in colour; for instance, Sirius and Rigel are bluish white, by comparison Capella looks yellowish white, while Aldebaran and Betelgeuse look red. The stars therefore do *not* all possess the same energy-curve and it follows that visual magnitudes do not suffice for the comparison of bolometric magnitudes. We have to say briefly how the latter are determined.

The apparent magnitudes of the stars can be compared photographically, i.e. by the relative intensities of their images on a photographic plate, as well as visually. Now suppose some star "A" is taken as the standard for both methods of comparison. Then if "S" is any other star, it is found in general that the apparent photographic magnitude of "S" relative to "A" differs from the apparent visual magnitude of "S" relative to "A". Since the maximum sensitivity of the photographic plate occurs at a different (higher) radiation-frequency from that of the human eye, it follows that the proportion of "photographic" to "visible" radiation is in general different for "S" and "A". This is merely another way of saying that the colours of "S" and "A" are different. But it is also evident that the difference between the photographic and visual magnitudes of "S" may be regarded as an actual *measure* of the colour of "S". As determined by a suitably standardized procedure, it constitutes in fact the accepted measure called the *colour-index*.

Now, were it known that, although the stars have different energy-curves, these are all black body curves and differ merely in the black body temperatures, then a theoretical colour-index could be calculated for each value of the temperature. Comparison of the results with the observed colour-index of any particular star "S" would then enable the black body temperature of "S" to be inferred. Since the ratio of total to visible radiation for that temperature could also be calculated, the total luminosity of "S" could also be inferred.

In other words, the assumption that the stars radiate like black bodies makes it possible to convert the visual magnitude into the bolometric magnitude when the colour-index is

known. This is the principle used in the practical determination of bolometric magnitudes. The effect of this allowance for colour is very appreciable; for example, Sirius is about 27 times as bright as the Sun in visible light, but according to this method the total luminosity of Sirius is about 39 times the total luminosity of the Sun. As an example of a less marked effect, the corresponding ratios for Capella are about 120 and 130, respectively.

The first consequence of all this discussion is merely that care must be taken always to specify what kind of luminosity is being considered; whether visual, photographic, or bolometric. The second is that, while visual and photographic magnitudes are derived directly from observation, the conversion to the more fundamental bolometric magnitude involves a specific hypothesis about the form of the energy-curve. Unfortunately, as we shall see, the hypothesis is not of general validity. It probably yields approximately correct results over a fairly wide range of colours but the accuracy in estimating true total luminosities is not very great.

This is a situation which calls for comment. It has been described at some length in order to illustrate, in a specific instance, the difficulties and uncertainties involved in securing the fundamental data of astrophysics.

The situation must not be misunderstood. The observations themselves are carried out with almost incredible care and precision. It is their *interpretation* to give the required fundamental quantities which is subject to considerable uncertainty in some cases. As a result of a most elaborate programme of observations a value may be got, for instance, for the total luminosity of a given star. Yet, when all is said and done, no one may care to say, in the particular case, whether the true value may not be perhaps twice or one-half the "observed" value!

This might appear to be a confession of gross uncertainty. Nevertheless, when we remember that the whole range of luminosities of the stars is of the order of 400,000,000 to 1, such a determination would place the given star in its correct position in the range with astonishing precision—a factor of two is neither here nor there! Therefore precise observations

those in which their uncertainties render comparison fruitless. Experience shows that this precaution leaves ample scope for trustworthy progress.

Effective temperatures. The value of the total luminosity of a star, together with the value of its radius, gives, as in the case of the Sun (p. 62), the value of the star's *effective temperature* T_e. This quantity really provides merely a convenient way of describing the total radiation per unit area of the star's surface. It is found that the effective temperatures of the stars range from less than 3,000 degrees up to 50,000 and possibly to about 100,000 degrees.

As defined, the effective temperature of a star is not a directly observable quantity. But the effective temperature is related to the colour, as is known both theoretically and observationally. Using a theoretical or empirical relation, the observed colour of a star may therefore be made to yield an *estimate* of its effective temperature (subject to some margin of uncertainty). Using this estimate, if the total luminosity of the star can be measured, we can then derive an estimate of its radius. As already mentioned, this principle provides the most common, but not the most accurate, means of determining stellar radii.

Continuous spectra. A great amount of observational work has been done on the energy-curves of the continuous spectra of the stars. In spite of the importance of the results, they cannot be brought within the compass of this book. All that can be said is that the curves differ to a very significant extent from black-body curves. The Sun, and stars of about the same effective temperature, are in fact somewhat exceptional in even the approximate resemblance of their energy-curves to black-body curves.

The interpretation of energy-curves belongs to the theory of stellar atmospheres. Where at present certain results are worked out on the assumption that the stars behave like black-bodies, they must in the future be worked out using a theory of stellar atmospheres in place of black-body theory. But, though the principles of the required theory are now well-understood it can scarcely be claimed to have been developed quantitatively with sufficient accuracy for this purpose.

Line Spectra: spectral sequence

A science which has to deal with a great number of differing individual objects cannot handle its material without an adequate system of classification. The characteristics of the stars which we have so far considered do not serve this purpose, partly because they cannot be determined with precision for any but a small fraction of the stars and partly because they are over-simple. Imagine a botanist attempting to classify plants according to their weight, size, colour, etc.! Just as the botanist looks deliberately for features that, in addition to exhibiting systematic variations, are of a sufficient complexity in their variations to make a reasonably detailed classification possible, so the astronomer must expect the classification of the stars to depend upon more complex characteristics than mass, size and luminosity.

Now the only observable feature of a star not yet mentioned is its line-spectrum. This is found to be in general certainly very complex and so it does hold promise of providing a means of detailed classification. Moreover, it has the fortunate merit of being by far the most readily observable intrinsic feature of the stars. Most important of all, it does vary systematically from star to star in the fashion about to be described.

Before entering upon this description it is, however, desirable to satisfy ourselves that the resulting classification is likely to be of fundamental and not just superficial significance. For we know from the case of the Sun that the spectrum is formed by an infinitesimal fraction of the total material of a star. Without going into any specific theory, we know, however,

is a very sensitive function of these three fundamental characteristics. It provides, in fact, a function of these characteristics which can be studied with incomparably greater accuracy than that with which any one of them can be studied by itself, hence its profound significance.

Classification. The spectra[1] of literally hundreds of thousands of stars have been carefully studied, classified and catalogued, and so the observational material available for the study is indeed vast.

The spectra of the majority of stars are in a general way similar to the Fraunhofer spectrum of the Sun. They show dark lines produced by the same chemical elements and, to a great extent, the self-same lines. But in the spectrum of any particular star it is in general noticeable that the lines occur with relative intensities different from those in the solar spectrum. Also, a certain proportion of stellar spectra show molecular, as opposed to atomic, absorption as more prominent features than does the solar spectrum. Finally, a small proportion of spectra contain bright, as opposed to dark, atomic lines.

Consider now a representative collection of stellar spectra. It is discovered by inspection that they can be arranged in a single sequence—we may imagine them all laid out in a single row—such that if attention be directed to any particular spectral line which figures prominently anywhere in the sequence, then, looking along the row from one spectrum to the next, the intensity of this line is seen gradually to rise to a maximum and then gradually to diminish again. Or, if in

[1] In the rest of the book, except when otherwise stated, we take "spectrum" to mean what corresponds to the Fraunhofer spectrum of the Sun, i.e. we disregard any general features of the continuous spectrum as such.

the case of a few lines the behaviour is something different, it is at any rate something perfectly orderly along the sequence. In fact, once the spectra have been sorted into this particular order, everything in them exhibits a nice smooth gradation along the sequence. Also it is found that, with relatively few exceptions, every stellar spectrum fits naturally into a place somewhere in the sequence. In other words, a spectrum can be labelled just by its position in the sequence.

This is a remarkable result. Looking at such a complex thing as a typical stellar spectrum we might have expected some much more elaborate system of classification to be required. It is as though the botanist had discovered some characteristic by which he could allot to every plant a definite position in one single "natural order", using the term natural order in its most literal sense!

The only way to introduce a classification into such a sequence is to make a convenient number of quite *arbitrary* divisions and to describe a member by the division in which it is located. The spectral sequence is divided into a number of spectral *classes*[1] six of which, called classes B, A, F, G, K, M, include about 99 per cent of known stellar spectra. The classes have been chosen so that a typical member of a particular class exhibits some important groups of spectral lines at their greatest relative prominence in the sequence. Each class is further divided into ten sub-divisions labelled 0 to 9. Thus Sirius has spectral class A0, Rigel B8, the Sun G2, Capella G4, Aldebaran K5, Betelgeuse M0.

Although there is no difference of opinion about the order in the sequence, different observatories have unfortunately placed the divisions at slightly different positions in the sequence. The correspondence between the various systems in use is, of course, well known.

Spectra near the beginning of the sequence are said to be "early" type and those towards the end "late" type spectra. These terms must not be used with any other connotation as,

[1] The terms and notation are partly the result of the historical development of the system. It may be mentioned that the sequence as a whole is called the Harvard sequence, and the classes are called Draper classes after the astronomer who first assigned them.

the sequence. The later Classes, from K5 onwards, show molecular bands to an increasing extent. Hydrogen lines are the only ones which appear throughout the sequence; they are strongest in Class Ao.

Class O, which precedes B and is one of the less abundant classes amongst known spectra, shows mainly bright lines. Certain late type spectra show some bright lines as well as absorption lines and bands. Otherwise bright lines occur elsewhere in the sequence only as exceptional features of some individual stars. Such features present problems of much interest but they need not distract us here.

Finally, it must now be stated that the foregoing actually describes only a first approximation to the true state of affairs. But it is such a good first approximation that it is legitimate to present it in the way we have done. The required slight elaboration will be mentioned in due course.

Interpretation. From purely spectroscopic data, without any theory of stellar atmospheres, it is readily seen that the progression in the spectral features along the sequence is purely one of steadily diminishing degrees of excitation and ionization of the atoms producing the spectra. This can be expressed by saying that *the spectral sequence is a temperature sequence* with the temperature decreasing from early to late type spectra. Moreover, the sequence as so far described cannot then be anything more than this; for were it to depend upon any other parameter independent of the temperature it would not be a *single* sequence.

The relevant "temperature" is really one defined by the state of excitation and ionization of the outermost layers of the stellar photosphere. The temperature so defined may not be numerically the same as the "effective temperature" of the star. But, without any detailed theory of stellar atmospheres, we may reasonably expect that, if one star is hotter than

another as judged by this "excitation temperature" it will also be hotter as judged by its effective temperature. That is to say, if the stars are arranged in order of decreasing effective temperature then they would necessarily be found to be in order of decreasing "excitation temperature", i.e. they would in fact be in the order of the spectral sequence.

This expectation is found to be borne out by the values of independently determined effective temperatures for stars of known spectral class (apart from the slight modification mentioned on p. 126). This means that a characteristic effective temperature can be associated with each spectral class. Therefore the determination of the spectral class provides also a determination of the effective temperature. This accounts for the importance of spectral class from the standpoint of general astrophysics.

Table II summarizes the correspondence between temperature and spectral class, giving the temperature only in round numbers.

TABLE II

Spectral Class	*Effective Temperature*
O	50,000–25,000
Bo	25,000
Ao	11,000
Fo	7,500
Go	6,000
Ko	5,000
Mo	3,500
N	3,000–2,000

CORRELATIONS

The fundamental observable characteristics of a star are most conveniently taken to be its mass M, absolute luminosity L, and spectral class S.

At this stage in our presentation of the subject there is no definite reason for expecting the quantities M, L, S to be themselves related in any way. But it is an obvious next step to examine the observational data to see if any such relation

varies with each of the other two, M, S. The results of doing this are of momentous importance.

In the first instance it would be simpler and more natural to work directly with the observed luminosity (visual or photographic) and spectral class, rather than with the derived bolometric luminosity and effective temperature, respectively. But, for the sake of subsequent theoretical discussion, it is far more convenient to use the latter quantities. Therefore, in what follows, the "observed luminosity" and the "observed effective temperature" mean the bolometric luminosity and the effective temperature derived from the observed luminosity and the observed spectral class in accordance with the best current estimates of the relations between these quantities.

This procedure does not vitiate the "observational" or "empirical" character of the correlations about to be described. All it does, so far as the work of the present chapter is concerned, is to determine the particular "code-numbers" in terms of which we choose to express the observed luminosities and spectral classes. It is only when, in the following chapter, we attempt to reach a theoretical explanation of the correlations that we need to regard these code-numbers as having the particular significance of bolometric luminosities and effective temperatures. At that stage we shall, in fact, so regard them—with the reservation that the numbers used may not be altogether accurate measures of these quantities on account of uncertainties in our knowledge of stellar spectral energy-curves.

Hertzsprung-Russell diagram. A diagram can be constructed in which the observed luminosities of the stars are plotted against their observed spectral classes; in such a diagram each "dot" represents the observed values of these characteristics for a particular star. The result is the famous Hertzsprung-Russell diagram, named after the two astronomers who first studied extensive sets of data in this manner.

The first figure shows a "skeleton" diagram with a few well-determined points. The second shows, in a schematic way, what the diagram looks like when a large sample of the available data has been plotted.

The Hertzsprung-Russell diagram reveals that luminosity and spectral class *are* related characteristics of the stars. It does more: it shows that the stars divide themselves into sets each characterized by a different form of the relation between luminosity and spectral class. The features of the diagram may be summarized as follows:

(*a*) The predominant feature is a set of stars represented by points along a track (m in Figs. 3, 4) which stretches obliquely right across the diagram. This is called the *main sequence* of stars. As we have seen, its existence is indicated by considerations regarding density and, as a matter of fact, it is found that stars which belong to the main sequence in the diagram would also be classified as main sequence stars on the density criterion.

(*b*) In the early spectral classes there is a set of stars of very small luminosity (w.d. in Figs. 3, 4). These are the *white dwarfs* and are the stars already mentioned as possessing phenomenally great densities. Their name accords with their being, in fact, white, in agreement with their early spectral class, and small, in agreement with their low luminosity. The figure contains too few points representing white dwarfs for any relationship between their characteristics to be revealed, beyond the fact that the points occur in one small part of the diagram.

(*c*) In spectral classes F and "later" there is a set of stars (g in Figs. 3, 4) of very much greater luminosity than the main sequence stars of the same classes. These are called *giant stars*; they are in fact very much larger as well as more luminous than main sequence stars of the same effective temperature. The giants form a fairly definite sequence of their own in which the variation of the luminosity from one class to another is quite different from that in the main sequence.

(*d*) In almost every spectral class there occur *super-giants* (s.g. in Fig. 4) of 10 to 100 times the luminosity of the normal

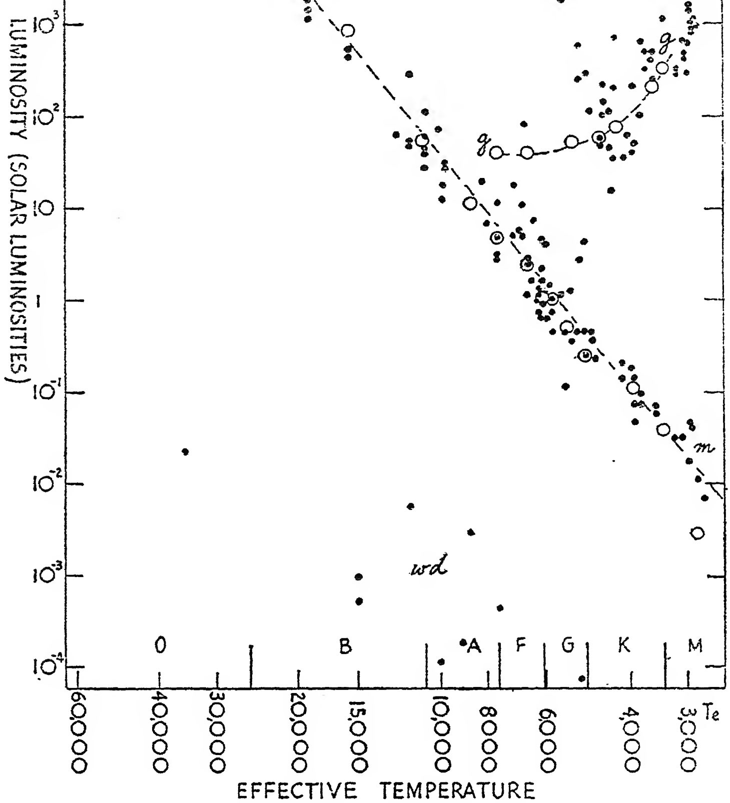


Fig. 3—Hertzsprung-Russell diagram. Dots represent data for individual stars. Open circles represent currently accepted mean values for various spectral classes. Spectral classes for main sequence are indicated at the foot and for giants at the top. *N.B.*—In this and the following diagrams the quantities are plotted on logarithmic scales.

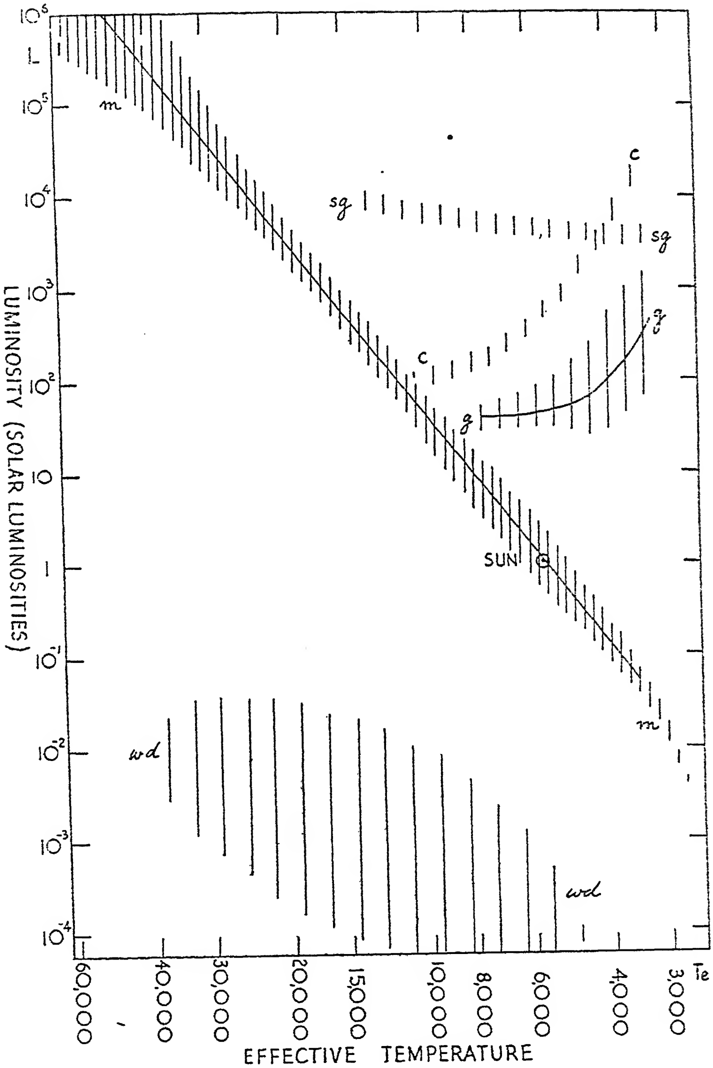


Fig. 4—Hertzsprung-Russell diagram—schematic. m = main sequence; wd = white dwarfs; g = giants; sg = super-giants; c = cepheid variables.

giants. The giants and super-giants are the stars previously mentioned as possessing very low densities.

Further analysis of the diagram is in fact possible, but the foregoing suffices for present purposes.

Relative numbers. The average giant star is over ten thousand times as luminous as the average of the known white dwarfs. So a giant has about the same apparent magnitude as has a dwarf at a distance one hundred times smaller; that is to say, the volume of space in which giants can be seen is of the order of a million times that in which white dwarfs can be seen down to the same apparent magnitude. Thus, on the ground of observability, a giant star has about a million times better chance than a white dwarf of being included in the diagram. Taking all considerations into account, the relative advantage of the giant may not be as large as a million to one, but it is certainly enormous. Consequently, the Russell-diagram as usually constructed does not, and is not intended to, give for the stars of the various sets a representation of the relative numbers in which they occur throughout the stellar system. To conclude that, just because the giants feature rather prominently in the diagram, they are a relatively numerous set of stars would be no more reasonable than to conclude, just because cabinet ministers and criminals feature prominently in the newspapers, that therefore cabinet ministers and criminals form a large proportion of the entire population.

A separate investigation is needed in order to estimate the relative numbers of the various sets of stars. It is sufficient to state here that the proportion of white dwarfs amongst all the stars in the Sun's neighbourhood in the galactic system has been estimated to be about 1 in 10. The proportion of giants is probably less than one per cent. Owing to their great luminosities, giants are in a majority amongst stars visible to the naked eye. But the great majority of *all* stars belong to the main sequence, while within the main sequence the proportion of stars of relatively small luminosity is known to be large.

Thus the main sequence forms literally the "main" bulk of the stars. In attempting to understand the constitution of the stars it is natural to begin with those of the main sequence.

And if, in the next chapter, we succeed in describing in pri ciple the solutions of the chief problems presented by the stars we shall have attained our main objective.

It must be finally pointed out that, if in the Russell diagra a curve is drawn to lie centrally along the track of the ma sequence, the points representing the observed L, S-valu for the individual stars are noticeably *scattered* about th curve. A certain amount of this scatter must be due to unce tainties in the observations. But, as regards order of magnitud it is believed to denote a real spread in the stellar characte istics, and we shall have to refer to it again. The same appli to the giant sequence. Also in each spectral class in whic giants occur, there are stars of intermediate brightness betwee giants and main sequence stars, but these intermediate typ are relatively rare and do not serve to destroy the distinctio which have been made between the two principal sets.

Absolute magnitude effect. What has been said about th spectral sequence being a single sequence is in agreement wit all the facts of observation down to all but the last detai We might indeed be somewhat surprised were the differenc between a giant and a main sequence star—a difference whic in class M is represented by a factor of as much as 300 in th linear dimensions—to show no trace of an effect upon th spectrum. Actually there is some effect which may be describe as follows: Any non-exceptional star can be assigned withou ambiguity to a particular spectral class. But within that clas the strengths of certain of the less prominent lines or bands i the spectrum are found to depend in a definite manner upo the absolute magnitude of the star. This *absolute magnitud effect* is so well-established empirically for classes F–M that i provides in practice one of the standard methods for obtainin absolute magnitudes of stars and hence their distances.

In classes F–M it is customary to denote the spectrum of typical main sequence[1] star by a prefix "d" and that of

[1] Main sequence stars of later spectral classes used to be calle "dwarfs", hence the choice of the prefix "d". But, except for thi purpose, the name has largely dropped out of use for stars of th main sequence.

typical giant by "g". Thus the solar spectrum is classified as dG2 and the spectrum of Capella as gG4.

It is found observationally that the brighter stars of a given spectral class have slightly lower effective temperatures than those of the less bright stars.

We may digress for a moment to notice the explanation of these effects. The appearance of the spectrum of a star depends upon the degrees of excitation and ionization of its atmosphere. These degrees depend upon the temperature and, in the case of ionization, upon the density also (see pages 50–1). But the density effect is subsidiary and is in any case such that to a first approximation a decrease in density has the same effect as a slight increase in temperature. It follows that, as we know, the spectral sequence is predominantly a temperature sequence. But it also follows that almost the same spectrum is produced by two stellar atmospheres of different densities if that having the lower density is at a slightly lower temperature. This explains why the effective temperature (or any other significant temperature) decreases slightly with increasing luminosity in any one spectral class, and thereby verifies that the atmosphere of stars of any such class have lower density the greater the luminosity of the star. Finally the density effect, though similar to a temperature effect, possesses small residual differences and these constitute the absolute magnitude effect as observed. These arguments, which are adequate to explain the principles of the observed effects, are as far as the matter can be taken without appeal to specific theories of stellar atmospheres.

Mass-luminosity relation. A diagram can next be constructed in which the observed luminosities of the stars are plotted against their observed masses. The result is shown in Figure 5. Such a diagram contains fewer entries than can be inserted in a Russell-diagram simply because the mass can be determined for far fewer stars than the spectral class. But there are ample data to show that there is in fact a close correlation between the two characteristics applying to all stars except apparently the white dwarfs. This is the *mass-luminosity relation*, first made famous by Eddington's theoretical studies in 1917.

That such a relationship should exist, as an empirical fact, must be truly astonishing to anyone making a first acquaintance with it. For it shows that the luminosity of a star depends

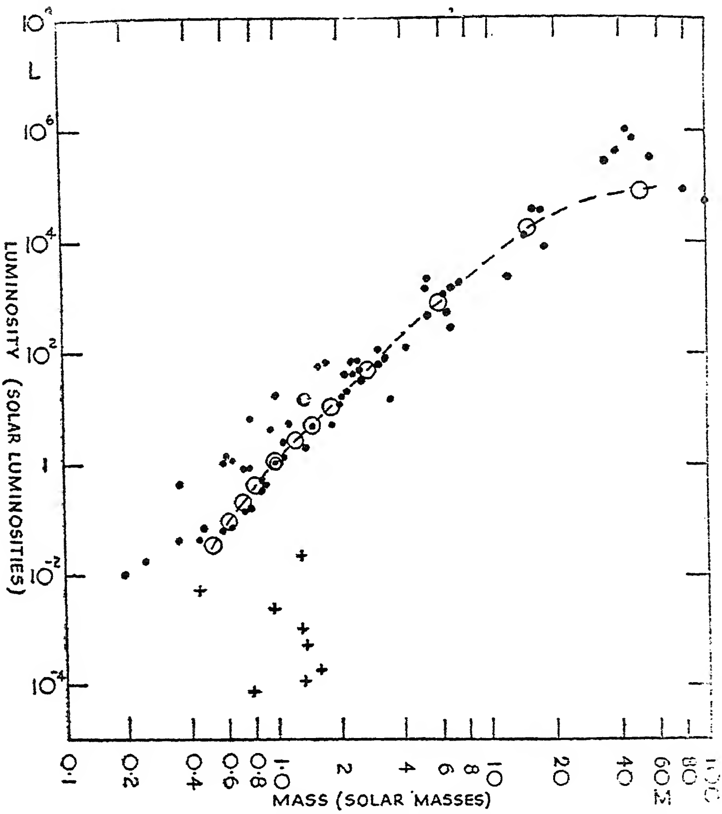


Fig. 5—Mass-luminosity relation; dots represent main sequence and giant stars; crosses represent white dwarfs; open circles represent average values for main-sequence stars.

primarily only upon its mass. General physical preconceptions would lead one to expect the luminosity to depend upon all kinds of features concerning the star's past and present physical state and chemical composition. But, even if it does, the

existence of the mass-luminosity relation shows that either the dependence must be relatively slight or else the features must themselves depend almost only upon the mass. Even so, having discovered the existence of main sequence stars, giants and super-giants, one might expect at least that any dependence of luminosity upon mass would be different for these different sets of stars. Yet this is not the case; the mass-luminosity relation evidently ignores or transcends such differences.

Such considerations almost compel the view that the mass-luminosity relation is of deeper significance than the class-luminosity relation. This is not surprising. For, if the luminosity, i.e. the energy-output, of a star depends principally upon the mass of a star it must do so by virtue of some very fundamental properties of matter. On the other hand, the spectral-class is practically the effective temperature, and for a given luminosity the effective temperature depends only on the radius. Therefore the class-luminosity relation can be regarded as showing that a star of given luminosity must assume a certain size. This must be a matter of stellar architecture and so is expected to depend upon other than the most fundamental properties of matter. What is surprising is that there should exist a mass-luminosity relation at all.

Chemical composition.—As in the case of the Russell-diagram, the mass-luminosity diagram shows a significant amount of scatter. The two diagrams together can be taken as showing that L and S both depend upon M. The scatter shows, however, that L and S do not depend upon M alone, but that the stars must differ from one another in regard to some other characteristic upon which L and S depend to a subsidiary extent. This characteristic cannot be any of those already studied, since we have already noted how these are related to L and S.

There is, however, a further regard in which it would be natural to expect one star to differ from another. This is its *chemical composition. A priori*, this might have been thought to constitute a major factor and some surprise may be felt that so little scope has been left for its operation. On a superficial

view one might have concluded, for instance, that the striking differences between various stellar spectra is simply evidence of differences of chemical composition. Actually, we have seen that the spectral differences are explained with no appeal whatever to differences of chemical composition. However, any differences in composition which could affect the luminosity must concern the interior of a star and therefore cannot be directly observable.

The empirical situation is that there is no known observable characteristic that accounts for the observed scatter in the *M–L* and *M–S* relations, but that the internal chemical composition is a characteristic which has not yet been allowed for. There is therefore strong presumption that it will be found to supply the required explanation.

It should be stated that there are apparently *systematic* differences between the Russell-diagrams for certain groups of stars in the sky. If these stars are all plotted in a single diagram, the differences merely contribute to the scatter we have mentioned. The scatter in the individual diagrams, on the other hand, is considerably less. Such effects are receiving much attention at the present time and are usually interpreted as indicating systematic differences in chemical composition.

CHAPTER IX

STELLAR CONSTITUTION

As in solar physics, there are two main branches of theoretical stellar astrophysics, that concerned with atmospheric constitution and that concerned with internal constitution. Both of these branches can be studied according to the same principles as those applied in the case of the Sun. An enormous amount of work has been done in this way. But to describe it here would for the most part introduce the reader to no fresh ideas. A new possibility does, however, enter at this stage, namely, the *comparative* study of the stars. This affords invaluable physical insight, chiefly regarding the internal constitution of the stars, and will form the main topic of the present chapter.

This may appear to be the hardest part of the book.[1] It is inevitable that a certain number of symbols should be employed. But we have to employ them in no more advanced manipulations than those expressed by the "laws of indices" which appear near the beginning of any elementary algebra book. The price to be paid is a small one for what will be found to be an understanding of nothing less than the general physical behaviour of the stars.

Method of similarity. We saw in the last chapter how the aggregate *observational* data on the stars reveal the existence of remarkable correlations between mass and luminosity and between mass and radius. We naturally ask whether it is possible to produce corresponding *theoretical* relations and to do so without detailed investigations for individual stars.

This question may prompt the reader to suggest our using the "method of dimensions" or, as it should be called more accurately, the "method of similarity". The reader may recall that this method can be used, for instance, to show how the

[1] The reader who wishes to go straight to the conclusions may turn to pages 150–1.

time of swing of a simple pendulum depends upon its length, without solving the equation of motion of a pendulum. He may expect to find that the method could also be employed to show how the luminosity of a star depends upon its mass, without solving the equations of stellar constitution.

The case of the pendulum is so simple that it gives the answer without any but the most general reference to the physics of the system. Consequently, the answer supplies very little new information! But in a more complex system, the method may give an answer only by applying it in some fairly detailed analysis of the physical processes which operate. The answer may then genuinely supply new insight into the interplay of the various processes. This we shall find to be the case in our present problems.

The procedure is familiar in many other contexts. It is that which can often be followed in engineering practice when it is necessary to predict how the performance of a particular type of structure will depend upon its size, the material of which it is made and, maybe, other factors such as its temperature.

Model star. We consider a theoretical model star constituted in the manner we are about to specify. We shall use the following symbols, some of which we have already introduced and others which we shall explain as we proceed:

M = mass
R = radius
L = luminosity, i.e. energy radiated per unit time.

At any point inside the star

r = distance from centre
d = material density
p = pressure
t = temperature
k = opacity per unit mass of the material
f = net outward energy-flux across unit area
e = rate of energy-generation per unit mass.

Quantities depending upon the chemical composition which will be assumed to be the same throughout the star

μ = mean molecular weight of the gas "particles"
κ = opacity constant of the material
ε = energy-generation constant of the material.

Thus capital letters denote characteristics of the whole star. Small letters describe conditions at any internal point and to describe the state of the star completely would be to say how these depend upon r from the centre ($r = 0$) to the surface ($r = R$). Greek letters denote chemical characteristics of the material. We shall also mention one or two universal physical constants which we shall denote by small capitals.

The constitution is assumed to be determined by the following requirements:

(a) The star is in mechanical equilibrium under its own gravitation. This means that the weight of the material above any level is supported by the pressure at that level.
(b) Radiation-pressure is negligible compared with material pressure throughout the star.
(c) The material everywhere behaves like a perfect gas. This means that[1]

$$p = \textsc{c}td/\mu \qquad (1)$$

where $\textsc{c}$ is the universal "gas constant".
(d) The star is in radiative equilibrium throughout. This means, in the first place, that the energy-transport is wholly radiative.

The rest of what it means can be understood by a physical analogy. Consider the case of fluid being forced through a slab

[1] Where the reader is unfamiliar with a particular formula such as this he can accept it as the expression of the stated physical requirement. Thus, to say that the material behaves like a perfect gas *means* that its pressure is directly proportional to the temperature and to the density, the constant of proportionality being a universal constant divided by the mean molecular weight of the gas. The formula merely puts this into symbols.

of porous material: the rate at which it comes through is directly proportional to the pressure-difference in the fluid between the two faces of the slab and inversely proportional to the obstructive power of the slab. In the same manner, when radiation forces its way through a star the flux is directly proportional to the radiation pressure-gradient, and inversely proportional to the obstructive power of the material. Now the radiation pressure is a universal constant times the fourth power of the temperature t. The obstructive power is measured by the density d times the opacity per unit mass k of the material; k is in fact defined by this property.

(e) The opacity is photo-electric. This is known to signify that it is given to a sufficient approximation by the formula

$$k = \kappa d/t^{7/2} \tag{2}$$

where κ is a constant depending upon the chemical composition of the material. The fact that the combined effect of all possible bound-free and free-free transitions in any gas, summed over all radiation-frequencies, depends in this way upon the density and temperature of the gas is a result we have to quote from quantum theory. It is known as "Kramers's law".

(f) The process of energy-generation is thermonuclear. This we shall take to mean that the rate of energy-generation is given to a sufficient approximation by

$$e = \varepsilon d t^{18} \tag{3}$$

where ε is a constant depending upon the chemical composition. This is Bethe's approximate expression of his result, the exponent 18 being that appropriate to conditions near the centre of the Sun when the significant process is the carbon-nitrogen cycle. For the sake of definiteness we adopt this particular value of the exponent.

Suppose we now tell a mathematician that a given mass M consists of material of given chemical composition and that it satisfies the requirements (a) to (f). He will find that, without any further assumptions, he can work out all its properties. He will conclude that the material is a spherical body in

which the density d, pressure p and temperature t have definite calculated values for every distance r from the centre and that the whole configuration has a calculated radius R and a calculated total energy-generation L.

The configuration deduced in this way from our requirements is our theoretical model star. It is a "possible" star in the sense that it is constructed in accordance with a physically consistent set of properties. Moreover, these are separately properties which are possessed by real matter. So we expect it to be a "useful" model. Whether or not it is matched by any actual star is a question we defer for the moment.

We are certainly not going to repeat the mathematician's calculations. We do not require even to quote his results. All we need is the assurance that these results *can* be got. Accepting only this, we find that we are ourselves in a position to discover how they must depend upon the assumed mass and chemical composition.

Family of Stars. Model stars can be constructed for a series of values of each of the quantities M, μ, κ, ε and in this way we can get a whole "family" of models. We can select any one member and regard it as the "standard" member with which we shall compare the others. Since we shall deal only with comparative values we can treat the values of M, μ, κ, ε, R, L for the standard as the unit values in which to measure each of these quantities. We shall then select any other member and call it for the moment simply "the star". So if we say the star has mass M, we imply that its mass is M times that of the standard, and so on for the other quantities.

The mass of the *whole* star is M times the mass of the standard and the radius of the *whole* star is R times the radius of the standard, while the chemical constants are, respectively, μ, κ, ε times those for the standard. We shall now see that the star can be supposed constructed from the standard by taking *every* distance in the star to be R times the corresponding distance in the standard, and the mass in *every* part of the star to be M times the mass in the corresponding part of the standard, while the chemical constants are everywhere multiplied by μ, κ, ε, respectively.

If this is so, the distance r or any point of the star from its centre is R times the distance of the corresponding point in the standard from the centre of the standard. We express this by the notation

$$r \sim R \tag{4}$$

Further, if lengths are multiplied by R, volumes are of course, multiplied by R^3. Also masses are multiplied by M. Therefore the mass per unit volume is multiplied by M/R^3, i.e. the density at any point of the star is M/R^3 times the density at the corresponding point in the standard. This is written

$$d \sim M/R^3. \tag{5}$$

We examine the requirements (a) to (f) in turn:

(a) If in the standard the mass of every part is multiplied by M, without any other change, then the weight of any part is multiplied by M^2. For the mass of that part itself is multiplied by M and the mass of every other part which exerts a gravitational attraction upon it is also multiplied by M.

On the other hand, if in the standard every distance is multiplied by R, then the weight of any part is multiplied by $1/R^2$. For every gravitational force has then to act at R times the original distance and we know that such a force varies according to the inverse-square law of distance.

Combining these factors, the weight of any part of the star is M^2/R^2 times the weight of the corresponding part of the standard. Therefore the pressure has to support this multiple of the original weight. But the area over which any pressure acts in the star is R^2 times the corresponding area in the standard. Therefore the pressure itself (meaning as always the pressure per unit area) is multiplied by $(M^2/R^2) \div R^2$. Thus we have

$$p \sim M^2/R^4. \tag{6}$$

(b) is assumed to hold good.

(c) Equation (1) asserts that the temperature is proportional to μ times the pressure divided by the density. Relations (5)

(e) We can take this before (d), for we now know from (5) and (7) how d and t are transformed. Using these results in equation (2) we get

$$k \sim \frac{\kappa(M/R^3)}{(\mu M/R)^{7/2}}$$

which reduces to

$$k \sim \frac{\kappa R^{1/2}}{\mu^{7/2}M^{5/2}}. \qquad (8)$$

(d) The radiation-pressure is proportional to t^4; equation (7) therefore shows that values at any two points in the star are each $(\mu M/R)^4$ times the values at the corresponding points in the standard. But the distance between these two points is R times the distance in the standard. Hence the radiation-pressure gradient, i.e. the variation of radiation pressure *per unit distance*, between the two points is multiplied by $(\mu M/R)^4 \div R$.

The obstructive power is, as we have said, $d \times k$. So, using (5) and (8), we see that it is transformed according to

$$d \times k \sim \frac{M}{R^3} \times \frac{\kappa R^{1/2}}{\mu^{7/2}M^{5/2}} = \frac{\kappa}{\mu^{7/2}M^{3/2}R^{5/2}}.$$

Since then the radiation flux is the radiation-pressure-gradient divided by the obstructive power, we have finally

$$f \sim (\mu^4 M^4/R^5)\,(\mu^{7/2}M^{3/2}R^{5/2}/\kappa)$$

which becomes

$$f \sim \frac{\mu^{15/2}M^{11/2}}{\kappa R^{5/2}}. \qquad (9)$$

(f) Using (5) and (7) in equations (3), we obtain

$$e \sim \left(\frac{M}{R^3}\right)\left(\frac{\mu M}{R}\right)^{18},$$

giving

$$e \sim \frac{\varepsilon\mu^{18}M^{19}}{R^{21}}. \tag{10}$$

Discussion. The meaning of these several results is that each of the separate requirements for the equilibrium of the star is satisfied if the quantities d, p, t, k, f, e at any point are those multiples of the same quantities at the corresponding point of the standard that are required by (5)–(10). Or we can interpret the relations by saying that the density, pressure, temperature in the star are determined by (5)–(7) in terms of those in the standard, that the opacity is then necessarily determined by (8), and that in the star thus obtained the flux and energy-generation must be as required by (9), (10) in terms of the standard.

The last two requirements, however, are not physically independent. For, if the star is in equilibrium, the flux must carry away energy at exactly the rate at which it is generated. So we must ensure that (9), (10) are consistent in this way.

Now the luminosity L is in fact the total flux across the whole surface of the star whose area is proportional to R^2. Therefore the boundary value of the flux is proportional to L/R^2. But the relation (9) must hold for the flux everywhere in the star and, in particular for that at the boundary. Hence (9) requires that,

$$L = \frac{\mu^{15/2}M^{11/2}}{\kappa R^{1/2}}. \tag{11}$$

We can at this stage restore the sign of equality. Were we to retain the sign $\sim$ it would mean that the ratio of the luminosity of the star to the standard is the ratio of the quantity on the right-hand side of (11) for the star to the same quantity for the

everywhere satisfy (10), and the whole mass is M times that of the standard. Hence (10) requires that

$$L = \frac{\varepsilon\mu^{18}M^{20}}{R^{21}}. \tag{12}$$

The condition for (9), (10) to be consistent is, therefore, the condition for (11), (12) to give the same value of the luminosity. Thus we must have

$$\frac{R^{21}}{\varepsilon\mu^{18}M^{20}} = \frac{\kappa R^{1/2}}{\mu^{15/2}M^{11/2}}$$

that is

$$R^{41} = \kappa^2\varepsilon^2\mu^{21}M^{29} \tag{13}$$

Thus the star is in equilibrium *if and only if its radius has this value depending upon its mass and chemical constants* (measured in terms of those for the standard).

Results. We can now list the properties of our family of theoretical model stars. It is evident from (13) that the exact values of the powers to which the various quantities occur in the formulæ may be rather awkward fractions. This is on account of the powers of t in formulæ (2), (3). These are in any case only approximate formulæ, so we do not attach significance to the exact values. The following results have

[1] Reverting to ordinary units and letting $M_\circ$, $R_\circ$, etc., apply to the standard, equation (11) would read

$$\frac{L}{L_\circ} = \left(\frac{\mu}{\mu_\circ}\right)^{15/2}\left(\frac{\kappa_\circ}{\kappa}\right)\left(\frac{M}{M_\circ}\right)^{11/2}\left(\frac{R_\circ}{R}\right)^{1/2}.$$

therefore been got by replacing awkward fractions by simple ones approximately equal to them; for instance 29/41 is replaced when convenient by 3/4.

Mass-radius relation. This is the relation (13) itself, but according to the simplification just described, it may be written with sufficient accuracy in the form

$$R = \kappa^{1/20}\varepsilon^{1/20}\mu^{1/2}M^{3/4}. \tag{14}$$

Mass-luminosity relation. If we substitute the value of R given by (13) in (11) or (12) we obtain, with sufficient accuracy,

$$\boxed{L = \frac{\mu^7 M^5}{\varepsilon^{1/40}\kappa}.} \tag{15}$$

Thus, for the stars of the family, the radius and luminosity are determined in terms of the mass and chemical composition. All the observable properties of the family depend upon these two relations. Though the remaining relations do not therefore express further independent properties, it will be found instructive to obtain them.

Radius-luminosity relation. Equation (13) may be regarded alternatively as giving M in terms of R; substituting in (11) or (12) we then have with sufficient accuracy

$$L = \frac{\mu^{7/2}R^7}{\varepsilon^{1/3}\kappa^{4/3}}. \tag{16}$$

Luminosity-effective temperature. The effective temperature T_e is defined so that the flux per unit area from the surface of the star is proportional to the fourth power of T_e. Thus the luminosity, which is the total flux from the whole surface, is given by

$$L = R^2 T_e^4, \tag{17}$$

remembering that all our quantities in such relations, including now T_e, are being measured in terms of those for the standard as unity.

We can use (16), (17) to express L in terms of T, instead of R, and we find

$$\boxed{L = \frac{\varepsilon^{1/7}\kappa^{1/2}}{\mu^{4/3}} T_e^{5\frac{1}{3}}} \tag{18}$$

Central temperature. If t_c is the central temperature of the star in terms of that of the standard, then equation (7) gives

$$t_c = \frac{\mu M}{R}. \tag{19}$$

Using equation (14) this becomes

$$t_c = \frac{\mu^{1/2}}{\kappa^{1/20}\varepsilon^{1/20}} M^{1/4}. \tag{20}$$

In passing, we should remark that a formula similar to (19) does not hold good for the effective temperature. A sufficient reason for this is that T_e is defined by the relation (17) and is not defined as the value of the temperature t at any point in the star.

Interpretation of results. We must still bear in mind that we are dealing with a family of theoretical model stars; we now discuss the properties of this family as revealed by our formulæ and compare them with the corresponding properties of actual stars.

The model stars exhibit a *mass-luminosity* relation expressed by equation (15), that is to say, for given chemical composition, the luminosity depends only on the mass. This much could indeed have been inferred on general grounds from the requirements (a)–(f) without any calculations. But the modest amount of calculation in which we have indulged tells us much more.

In the first place equation (15) shows that *the luminosity is proportional to the fifth power of the mass.* Fig. 6 is a reproduction of part of Fig. 5 with the addition of the line showing this law; Fig. 6 shows only the most reliably determined points of Fig. 5. We see that it reproduces the empirical mass-luminosity law not too badly except for the largest masses. Actually the

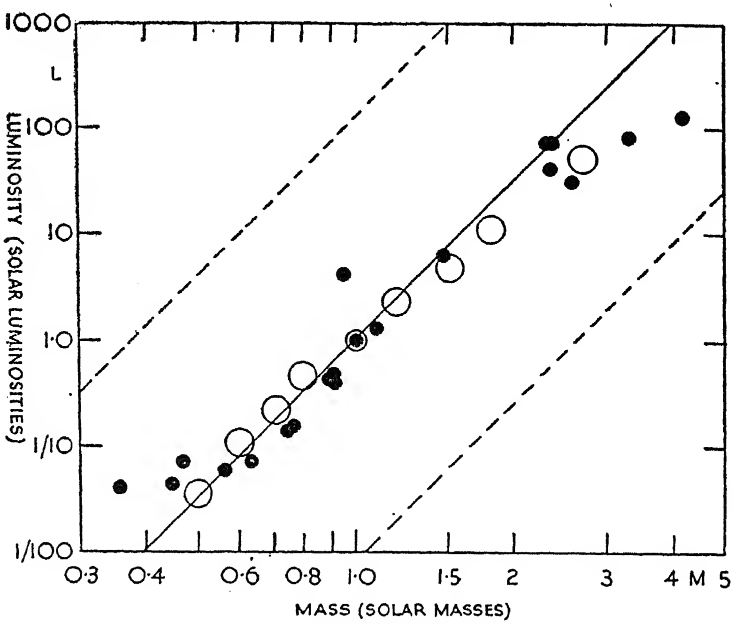


Fig. 6—Mass-luminosity relation: lines are drawn for luminosity proportional to fifth power of mass. The interval between the dotted lines is that permitted by differences of chemical composition. Open circles represent average values for main-sequence stars.

empirical law is (approximately) $L \propto M^{4\frac{1}{2}}$, for masses less than about half the solar mass, changing over to $L \propto M^{3\frac{1}{2}}$ for masses exceeding about one-and-a-half solar masses.

In the second place, the theoretical relation (15) exhibits a quite astonishing feature in regard to the dependence on chemical composition. The constant ε of the law of energy-generation occurs only to the power $1/40$. This means that

the luminosity of the star is excessively insensitive to the strength of the energy-sources, and not, as we might suppose, directly proportional to this strength. The formula would mean that, if the chemical composition changes in such a way that the strength of the energy-sources is diminished by a factor of, say, 100, then the resulting change in the luminosity would be only about 12 per cent. Were the formula to apply to the Sun, it would mean that the Sun would go on shining with a scarcely perceptible change in its present brightness, so far as this depends upon its energy-sources, for many times the duration of its whole past life. Moreover, owing to the fact that the factor occurs in the denominator in (15), a decrease in the strength of the sources would produce a slight *increase* in the luminosity!

The explanation of this seemingly unnatural behaviour of the star resides in the form of the energy-generation formula (3). Whereas the potential strength of the sources as determined by the chemical composition is represented by the factor ε, the rate at which the sources generate energy is seen to depend also upon the density and the temperature. Owing to the high power to which the temperature occurs, the temperature dependance is predominant.

Since this is a feature of fundamental importance in understanding the physics of the stars, we may follow it through in detail. Suppose the energy-sources in a star do suffer a sudden slight weakening. Then the rate of energy-supply ceases to be adequate to keep the star distended at its existing radius. Therefore it must contract. This is verified by equation (14), which shows that a decrease in ε will produce a slight corresponding decrease in R. But a slight contraction will produce, according to the relation (7) a slight increase in the temperature inside the star. Owing to the high power of the temperature in the energy-generation law, this produces a relatively large increase in the rate of energy-generation. So the net result is that the star fetches up with a very slightly smaller radius and a very slightly increased luminosity.

These arguments are based in the first instance upon the model of a star which we are using. But they can be taken as showing quite generally that, if the rate at which energy-

sources operate is very sensitive to temperature, then the luminosity of any star must be very insensitive to the "strength" of the sources. Conversely, they show that very little can be inferred about the "strength" of the sources by comparing the theoretical and observed luminosities of a particular star.

Having inferred that the chemical composition has little effect upon the luminosity of a star in so far as energy-generation is concerned, we may now consider the other factors in equation (15) which depend upon the composition. Take first the factor μ^7. Under the conditions of the Sun's interior, we saw (pages 87–8) that μ must lie[1] between $\frac{1}{2}$ and about 2. The same arguments would apply to the stars in general. If then in figure 6 the line showing the theoretical mass-luminosity law is supposed to apply to the value $\mu = 1$, the lines for $\mu = \frac{1}{2}$ and $\mu = 2$ would lie in the positions indicated, and the strip between them would show the region within which L can vary in so far as its dependence upon μ is concerned. It is seen that the latitude permitted to L for any given mass is ample to cover the scatter of the observed values. We conclude therefore that the scatter in the mass-luminosity diagram is consistent with an explanation depending upon differences in chemical composition, as far as our model is applicable. Since this conclusion has been reached without considering also the effect of possible variations in the value of κ, there is now no need to consider these in regard to this particular question.

The model stars exhibit also a *luminosity-effective temperature* relation expressed by equation (18). Recalling the connection between effective temperature and spectral class (page 120) we see that this is a theoretical relation between luminosity and spectral class. In other words, it gives the *Hertzsprung-Russell diagram* for the model stars.

Equation (18) shows that the *luminosity is proportional to the power* $5\frac{1}{2}$ *of the effective temperature*. Fig. 7 is a reproduction of part of Fig. 3 with the addition of the line showing this law. We see that it reproduces the general run of the

[1] Here we refer to the value in terms of the mass of the hydrogen atom. But we are concerned only with relative values in the present argument and the unit employed is irrelevant.

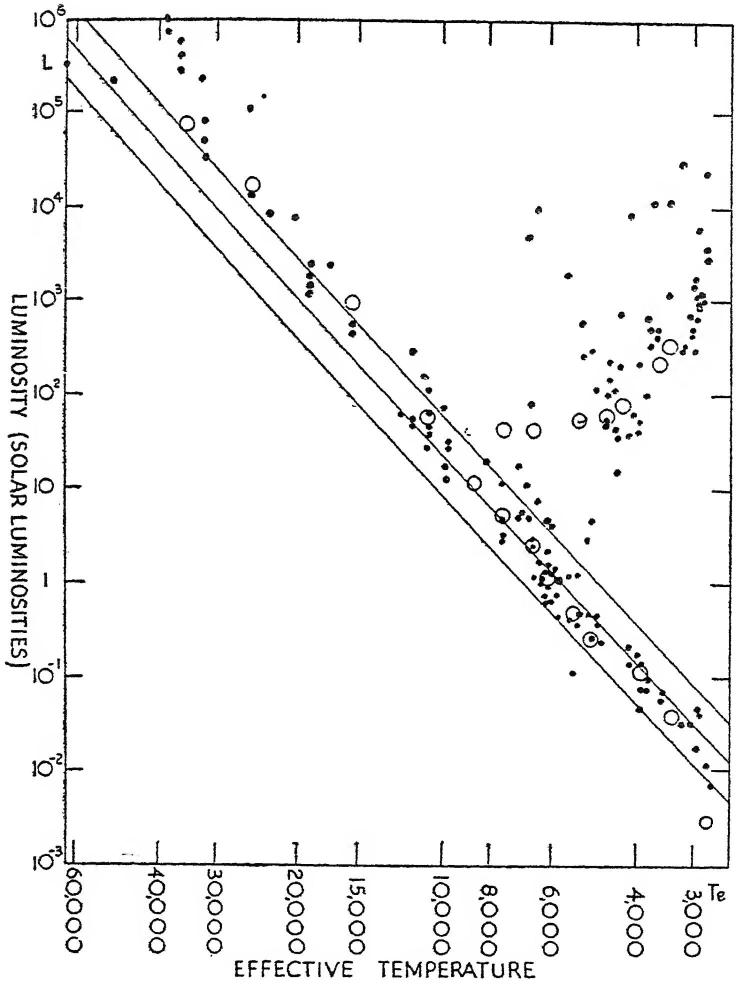


Fig. 7—Hertzsprung-Russell diagram: lines are drawn for luminosity proportional to the power $5\frac{1}{2}$ of effective temperature. The interval between the outside lines is that permitted by differences of chemical composition (mean molecular weight).

empirical law in a highly satisfactory manner. This conclusion is restricted to the *main sequence*, but, in view of the predominant rôle of this sequence, this is what we should wish if our models are to provide general insight into the behaviour of actual stars.

We then note further that the significant range of values of μ would give the strip indicated in Fig. 7, on the arbitrary assumption that the first line has been drawn for $\mu = 1$. The width of this strip is again of the right order of magnitude to suggest that the scatter in the Hertzsprung-Russell diagram is consistent with variations in chemical composition.

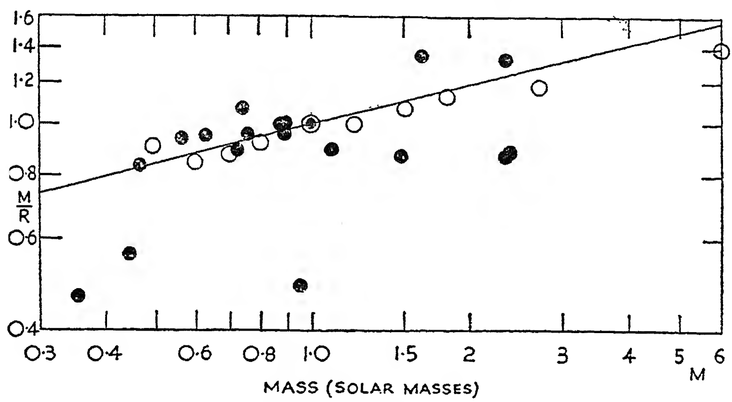
Finally, we see from equation (19) that, for given chemical composition, the central temperature of a model star is proportional to the ratio M/R. Looking back to its derivation, we see that this conclusion depends only upon the assumptions that the model is in mechanical equilibrium and that its material behaves like a gas whose pressure predominates over radiation pressure. It depends upon no assumptions about energy-generation or energy-transport. It depends in fact upon such few and elementary assumptions that we can almost say that, if values of M/R are got from observation for actual stars, then these *are* the values of the central temperature in terms of that of any one of them.

Fig. 8 shows the observed value of M/R for a number of the best determinations. We see that in the whole of the main sequence the value of their ratio varies by little more than a factor 2. We are driven to conclude that the central temperature varies only to this very small extent even though for the same stars the luminosity varies by a factor of over a million.

This remarkable near-constancy of the central temperatures of the main sequence stars was inferred by Eddington many years before its full significance was appreciated. Its great importance is that it points conclusively to a process of energy-generation which depends upon a very high power of the temperature. Thus, in order to get a million-fold increase in the rate of energy-generation for a two-fold increase in the temperature (disregarding, for the moment, the effect of other factors) this rate would have to be proportional to about the 20th power of the temperature. Now this is in fact

ɔut the power given by Bethe's law. This argument is the ɔst convincing one for the applicability of Bethe's law to ː main-sequence stars; it carries much more weight than the ːeement of the predicted luminosity with observation for y particular star.

The argument just given avoids *assumptions* about energy-neration and energy-transport by making use of observed lues of both M and R. If we now make the assumptions on



g. 8—Mass-radius or central temperature relation; line is drawn r M/R proportional to $M^{1/4}$. Open circles show mean values for main-sequence stars.

hich our model is based we obtain equation (20) expressing e central temperature in terms of the mass and chemical ɪmposition alone. It shows that the central temperature is sensitive to the chemical composition and varies as $M^{1/4}$. his variation is shown in Fig. 8 and is seen to agree well ith the interpretation of the ratio M/R as a measure of the ːntral temperature. This is the most instructive way of ːhibiting the *mass-radius relation.*

alidity of the model. We now come to the crucial question of ɩe validity of our model as a reproduction of an actual star.

All the properties we have assumed are properties of actual ɩatter and our star could exist as a real physical body. We

should find, however, from a more detailed mathematical treatment that conditions near the centre are such that convection would tend to be set up. Though the model would be in equilibrium, its equilibrium would be unstable in this regard and so would be a type of equilibrium which would not be attained in actuality. As we know from the case of the Sun (page 105) an actual star would possess a central core in convective equilibrium.

Apart from this, our model would not differ fundamentally from the constitution which the Sun is believed to possess. Also the convective region contains, in the case of the Sun, only a rather small fraction (about 12 per cent) of the total mass. Moreover, in this region the distribution of density and temperature are not grossly different from what we should infer by assuming radiative equilibrium throughout. So we can take it that our model, for a mass equal to the Sun's mass, would be a fairly good rough approximation to the actual Sun.

Further, however, a model may be only a rough reproduction of the actual system it represents but may nevertheless show perfectly reliably how the various characteristics are related amongst themselves. In our case, we expect, on general physical grounds, that our models are quite near enough to actual stars to show reasonably well how, for instance, the luminosity varies from one star to another when the mass varies. This is what we expect and this is what we have in fact found from the comparisons with observation discussed in the last section. Having found this, we may then have considerable confidence in the physical interpretation of the observations provided by the model.

Actually, however, the situation is even better than this. For if, instead of the model which we have been using, we take one that does in fact possess a convective core, then it can be shown that *all the preceding results still hold good.* More precisely, if the requirements stated on pp. 133–4 are relaxed so that (d) and (e) are required to apply only to the outer zone in radiative equilibrium, then the results are the same as before and it is only the derivation that has to be modified. (We should, of course, obtain a different "standard",

but here we are concerned only with the formulæ for *comparing* the characteristics of any member of the "family" with those of the one selected as "standard". These *comparative* formulæ are not affected by the generalization which takes account of the convective core.)

There are two other respects, besides the possibility of convection, in which the model may prove to be deficient. Firstly, it is constructed on the supposition that radiation pressure may be neglected so far as the mechanical equilibrium is concerned. We might find that the model turned out to be inconsistent with this supposition. But we know that it is justified in the case of the Sun and we could actually use our own results to infer that it must therefore hold for all main sequence stars of not too large a mass.

Secondly, the model is constructed on the supposition that photo-electric opacity predominates over electron opacity. Again it might be found that the inferred physical conditions would contradict this supposition. But it is believed that in the case of the Sun the two sorts of opacity are of about the same order of magnitude. Our results could be used to infer that for masses less than the solar mass photo-electric opacity must then in fact predominate while for appreciably larger masses electron opacity may be the more important.

We do not pursue these elaborations, but it is probable that they account for some of the discrepancy between the simple theory and the observations in Figs. 6, 7.

Conclusions. The conclusions of all this discussion can probably be most simply stated in the following terms; they apply to stars of the main sequence:

> All the stars generate their energy by a thermonuclear process whose rate is proportional to a high power (about 20) of the temperature. This process is almost certainly the carbon-nitrogen cycle.
>
> This leads to a dependence of the stellar characteristics upon the stellar mass in agreement with the empirical results revealed in the mass-luminosity relation, the Russell-diagram, and the mass-radius relation.

The scatter in the empirical relations can be accounted for by variation in chemical composition.

The direct effect of the chemical composition upon the rate of energy-generation is, however, of minor importance. Until the raw material of the process is almost entirely consumed, the luminosity of a star will show relatively little change.

The brevity of this statement must not obscure the magnitude of its significance. On the astronomical side, the empirical relations are the results of incredible labours in amassing and systematizing observational data. On the physical side, the explanations depend upon a comparable aggregate of laboratory data and upon the whole structure of physical theory founded upon them. In addition, the establishing of the "astrophysical" relationships has required a vast amount of mathematical investigation on the part of theoretical astrophysicists. The achievement, at this stage, is a broad understanding of the constitution of some 90 per cent of all the stars.

White Dwarfs

As judged by abundance amongst the stars, the most important set that do not conform to the results we have been discussing are the white dwarf stars. We recall that these have enormous mean densities and abnormally small luminosities as compared with stars of similar masses belonging to the main sequence. The divergence from the results for the main sequence is clear from the positions occupied by the white dwarfs in Figs. 4 and 5 and is seen to be very drastic in character.

All other known stars (except for certain sorts which are believed to be on the way to becoming white dwarfs) conform more or less closely to the mass-luminosity relation. Indeed, the foregoing discussion makes it appear rather difficult for a star *not* to conform! It is, therefore, the small luminosity that is so surprising a feature of the white dwarfs, rather than the stupendous density. Once we have accustomed ourselves to the idea that bare nuclei and free electrons can be packed

together into a vastly smaller amount of space than can the corresponding entire atoms, there is no further cause for surprise at the existence of matter possessing white-dwarf densities.

Nevertheless, if an attempt is made to apply to white dwarfs the theory of stellar structure which is successful for other stars, it is found that the results do not merely disagree with the observed characteristics but that they present an unexpected paradox.

This paradox was first encountered by Eddington, and the following was the train of thought leading to it. Were we to imagine a star to exhaust its supply of sub-atomic energy, we should in accordance with our ordinary physical ideas expect that it would gradually cool down and finally attain the state of a cold solid body. Apart from questions as to what the intermediate stages might be, there would appear to be no difficulty in envisaging this as a possible[1] final state for, say, the Sun. A cold solid of the same mass would certainly possess less energy and, so far as energetic considerations are concerned, the Sun could attain that state by losing energy, i.e. it could grow cold by the simple process of cooling.

A white dwarf presents quite a different case. Its density is about a million times greater than that of any material in the solid phase. Therefore, in order to solidify, it must expand. Such expansion against the mutual gravitation of its parts requires a great increase in gravitational energy. Eddington's calculations, made along similar lines to those which he had found to be successful for other stars, led him to the result that the heat energy of a white dwarf would be inadequate to supply the needed gravitational energy. Thus, divested of sources of sub-atomic energy, a white dwarf has already less total energy than a completely cold body of the same mass. As Eddington says, it would be a system "continually losing heat, but with insufficient energy to grow cold"!

It therefore follows that some conceptions which are

[1] Actually it is not a possible state, as will appear in due course, but we are here describing the considerations which led to the paradox.

ordinarily valid have to be modified in order to deal with matter at the densities found in white dwarfs in regard either to its present or its ultimate physical state or both. Actually, the modification is not far to seek. Eddington's paradox is only of historical interest; but it has been mentioned in order to show that the quantum theoretical considerations about to be introduced constitute not merely an academic refinement but provide an escape from physical absurdity.

Dense matter. Consider first the material in a typical region inside the *Sun* and suppose, say, one litre of this material to be "boxed up" in some sort of rigid enclosure. The enclosed material will consist of a large number of atoms in a highly ionized state, i.e. a number of nuclei, ions and free electrons, and these will be in violent thermal agitation. Suppose then that all the heat energy is somehow abstracted so that the interior of the enclosure is reduced to zero absolute temperature. The nuclei, ions and electrons will then be re-associated into entire atoms. At the density which the material possesses, these atoms will be rather tightly packed, since they are so much larger than the particles into which they had been dissociated. But there may be just room enough for them in the litre volume. The material would, of course, no longer be anything like a perfect gas but would be a cold solid or something like it.

Further, in this imagined cold state, the electronic systems of the reconstituted atoms would all be in their ground states. It should be noted that the individual electrons would then possess the orbital motions proper to those states. Thus the electrons are not reduced to rest at zero temperature, but the orbital motions do not count as thermal motions.

If, finally, the enclosed material is re-heated it will recover its original state and in particular its original degree of ionization. It is therefore natural to speak of such ionization, which can in this way be brought about by heating, as *thermal ionization*.

Suppose next that similar treatment be applied to the material in a typical region inside a *white dwarf*. Thus we imagine the material inside a litre volume to be reduced to

zero temperature. It is seen at once that this cannot now result in the reconstitution of entire atoms. For the material is about a million times as dense as in the previous case. Since in that case the reconstituted atoms could only just be packed into the litre, in the present case they would occupy about a million litres even if they were packed as tightly as possible. In the postulated rigid enclosure of one litre volume it is therefore impossible to achieve anything remotely like an assembly of normal atoms. That is to say, we may not think of electrons as becoming attached to particular nuclei, for if they did the resulting systems would have atomic dimensions and there would be no room for them. The only alternative is to regard the whole of the enclosed material as forming one single system of nuclei and electrons.

Simple electrostatic considerations indicate the general nature of the system. The positively charged nuclei must distribute themselves through the volume in some uniform manner and the negatively charged electrons must distribute themselves through the intervening space. The system must in fact be somewhat like a crystal lattice, except that all the electrons must be regarded as being shared between the nuclei since there is no room for any of them to describe orbits around individual nuclei. Consequently it is appropriate to speak of the aggregate of electrons in the enclosure as the electron "gas".

Quantum theoretical considerations (of the sort discussed in Chapter III) show further that the system must be quantized. In the supposed state of zero temperature, it is sufficient to regard the nuclei as remaining fixed and to assert that any electron in the system must occupy a quantum state or quantum orbit in the combined field of the nuclei and other electrons, just as in a single atomic system an electron occupies a quantum state in the field of the single nucleus and of the other electrons. At zero temperature the whole system must be in its "ground state". But once again it must be noted that the individual electrons would possess the motions appropriate to this state. These motions are again not thermal in character, but as before there is no question of the electrons being reduced to rest even at zero temperature.

Degenerate gas. The electron gas in the ground state which we are considering is said to be *degenerate*—a term connected with the more technical quantum theoretical treatment and not intended to be expressive of the physical ideas.

The comparison of the degenerate electron gas with the electronic system of an atom may be pursued further. If the latter is considered as being built up electron by electron, it is known that the successive electrons are less and less tightly bound to the system, i.e. they go into quantum states of increasing energy. Thus the average energy per electron increases with the increasing number of electrons. The same must be true of the degenerate gas; if more electrons are put in, i.e. if the electron density is increased, the average energy per electron is increased.

Now the electron density could be increased by the simple process of compressing the gas. If this increases the average energy per electron, it necessarily increases the total energy in the enclosure. Therefore, to supply this energy, it is necessary to expend mechanical work in order to reduce the volume of the enclosure. That is to say, despite the hypothesis that its temperature is zero, the degenerate gas must exert a definite pressure. It follows also that this pressure depends only on the electron density ρ (say). The pressure can be calculated by quantum mechanics and is found[1] to be proportional to $\rho^{5/3}$.

The volume of the enclosure being kept fixed, suppose that the temperature is raised from zero to some value T. The enclosed electron gas is still regarded as a quantized system. Then, for a quantized system, the increase in temperature is shown by the system passing from its ground state into some excited state. The calculation of the pressure of the electron gas can then be repeated for its excited state.

The result is somewhat surprising. For the ranges of density and temperature which subsequently turn out to be relevant to conditions in white dwarfs it is found that the pressure is

[1] It is generally considered that, under certain conditions, relativity effects have to be taken into account; under these conditions the formula is modified. But it would be outside our scope to discuss such complications.

Also it will be observed that the effect of varying the number of nuclei per unit volume has been ignored; this can be justified.

scarcely different from that of the completely degenerate gas, i.e. for $T = 0$. Thus the "zero-point pressure", as we may call it, accounts for effectively the whole pressure in the interiors of white dwarfs; so this pressure can be taken to depend upon the density only in the manner mentioned above.

Lastly, since the electrons in the material are not bound to particular nuclei even when the temperature is supposed reduced to zero, the material must be regarded as completely ionized even at zero temperature. Now, ionization at zero temperature is certainly not "thermal" ionization. It is due, as we saw, to the great density or, what comes to the same thing, to the great pressure. The ionization in the present case is therefore described as *pressure ionization*.

At this point the reader may enquire why quantum mechanical considerations, which are found to be so essential in the present case, were ignored in dealing with free electrons in ordinary stars. The fact is that we have to do with extreme cases. The quantum mechanical treatment *could* be formulated so as to cover all possible cases. But, for the order of density occurring in ordinary stars the results would be indistinguishable from those got by treating the electrons as a classical ideal gas. For the densities in white dwarfs, on the other hand, the results are indistinguishable from those got by treating the electrons as a completely degenerate gas in the quantum sense.

Constitution. We may consider first the very well known white dwarf, the companion of Sirius (known as Sirius B). The observational data for this star, stated in round numbers, are:

Mass . . .	1 solar mass
Radius . . .	1/50 solar radius
Mean density . .	125,000 solar mean density = 175 kg./c.c.
Surface gravity . .	2,500 solar gravity = 70,000 Earth's surface gravity
Spectral class . .	F0
Effective temperature .	9,500 degrees

For a reason which will shortly appear, let us consider the outermost $\frac{1}{4}$ per cent of the total mass. From the above figures

atmospheres may be applied in order to calculate the distribution of density and temperature. If the whole of the layer just mentioned may be treated in this way, then, as S. Chandrasekhar finds,[1] the density at the base of the layer would be about 2 kilogrammes per cubic centimetre and the temperature about 20 million degrees.

The importance of this result is that, according to the formulæ of quantum theory, matter of this density and temperature would be effectively completely degenerate. It also follows, by a slight elaboration of the argument, that the matter at any deeper level must *a fortiori* be in that state. Consequently, less than $\frac{1}{4}$ per cent of the matter in Sirius B is non-degenerate.

Chandrasekhar and others have shown that a similar conclusion holds good for all known white dwarfs. It follows that, apart from a layer which is only skin-deep, a white dwarf is a configuration of wholly degenerate matter.

Now it is possible, for any mass M of material of any assumed chemical composition, to calculate the equilibrium configuration when the material is postulated to be wholly degenerate. Indeed the calculation is a relatively simple one since the pressure of degenerate matter depends only on the density, so that the problem of mechanical equilibrium can be solved independently of questions of temperature distribution and energy-transport.

The solution of the problem gives a definite theoretical value of the radius R for each value of M and each assumed chemical composition. It can be asserted that the theoretical value of the radius is in satisfactory agreement with the observed value, for a plausible chemical composition, in the cases where a comparison is possible.

[1] His calculations employ, naturally, the actual observational data and not those quoted in round numbers. The latter are used here because they make the orders of magnitude more obvious.

degenerate, they compute radii which are in satisfactory agreement with the observational data.

The conclusion that *the white dwarfs consist of degenerate matter* is therefore extremely well established.

This result is particularly interesting because, although we are accustomed to the idea that properties of atoms can be understood only with the aid of the quantum theory, we normally suppose that the properties of gross matter can be understood without its explicit use. But the white dwarfs in the aggregate contain probably something like one-tenth of all stellar matter and their matter is now found to possess gross properties which cannot be explained except with the aid of the quantum theory.

Luminosity. If we have a quantity of material of the same mass as, say, the Sun, and if it contains potential sources of thermonuclear energy-generation like those of the Sun, then those energy-sources must operate and the material must adopt a configuration and produce a luminosity like those of the Sun. In short, the material must *be* the Sun or a star very like the Sun. This is the inescapable conclusion of all accepted views regarding stars of the main sequence.

Now a white dwarf consists of a quantity of material of mass comparable to that of the Sun. It follows that this material cannot contain potential energy-sources comparable to those in solar material. Otherwise the material would perforce form a star comparable in luminosity to the Sun.

In other words, the fundamental thing about a white dwarf must be its small luminosity and not its high density. The small luminosity must result from a real deficiency of sources of energy-generation. This deficiency is what hinders its behaving as a main-sequence star; its high density must then in some way result from this deficiency.

The theory of thermonuclear release of energy shows that

degenerate, they compute radii which are in satisfactory agreement with the observational data.

The conclusion that *the white dwarfs consist of degenerate matter* is therefore extremely well established.

This result is particularly interesting because, although we are accustomed to the idea that properties of atoms can be understood only with the aid of the quantum theory, we normally suppose that the properties of gross matter can be understood without its explicit use. But the white dwarfs in the aggregate contain probably something like one-tenth of all stellar matter and their matter is now found to possess gross properties which cannot be explained except with the aid of the quantum theory.

Luminosity. If we have a quantity of material of the same mass as, say, the Sun, and if it contains potential sources of thermonuclear energy-generation like those of the Sun, then those energy-sources must operate and the material must adopt a configuration and produce a luminosity like those of the Sun. In short, the material must *be* the Sun or a star very like the Sun. This is the inescapable conclusion of all accepted views regarding stars of the main sequence.

Now a white dwarf consists of a quantity of material of mass comparable to that of the Sun. It follows that this material cannot contain potential energy-sources comparable to those in solar material. Otherwise the material would perforce form a star comparable in luminosity to the Sun.

In other words, the fundamental thing about a white dwarf must be its small luminosity and not its high density. The small luminosity must result from a real deficiency of sources of energy-generation. This deficiency is what hinders its behaving as a main-sequence star; its high density must then in some way result from this deficiency.

The theory of thermonuclear release of energy shows that

carbon-nitrogen cycle using such traces of hydrogen as still remain unconsumed. The degeneracy of the material is mainly the affair of the electrons and will have no appreciable effect upon the nuclei and their participation in nuclear reactions. But the degeneracy has another consequence which has not yet been mentioned; it renders the material highly transparent to temperature radiation for the range of temperatures which we expect to find in a stellar interior. It follows that only a relatively small temperature-gradient is needed to push the radiation-flux through the material. So the interior of a white dwarf must be approximately isothermal. But we know also that, almost irrespective of the concentration of the reagents, the carbon-nitrogen cycle demands a temperature of the order of 10 million degrees in order to produce any observable luminosity. We can therefore infer that the temperature is of this order throughout the interior of a white dwarf.

This is confirmed by the fact that calculations such as those of Chandrasekhar give a temperature of this order at the level where, proceeding inwards from the surface, degeneracy first sets in.

Conclusions. For the reasons which have been stated we conclude that the white dwarfs consist almost entirely of degenerate matter, that they retain only a trace of hydrogen in their interiors and that the lack of hydrogen accounts for their small luminosities, that their luminosities are produced by the carbon-nitrogen cycle operating upon the trace of hydrogen present, and that the internal temperature is of the order of magnitude of 10 million degrees throughout. The discussion has shown that this set of conclusions is entirely self-consistent.

These in turn lead almost inevitably to the further conclusion that every ordinary star must ultimately become a white dwarf as a result of the burning up of its hydrogen. This, roughly speaking, is what is generally believed to be the case and some of the possible processes of this evolution will be considered in the next chapter.

As to what happens once a star reaches the white dwarf state, we see in the first place that no further contraction is possible. For the radius depends only on the mass and not on

the internal temperature so that further cooling does not affect the radius. We see, incidentally, that a white dwarf cannot release any gravitational energy.

In the second place, the very small luminosity of a white dwarf shows that whatever sources of thermonuclear energy-generation it possesses are being used up exceedingly slowly. But they will ultimately become exhausted. After that, so far as we know, the star, which will then be invisible, must cool down to the absolute zero of temperature. The material will remain completely degenerate.

Russell and also D. S. Kothari have shown that a body of stellar mass cannot become solid in the classical sense even at zero temperature. But Eddington's paradox is no longer a difficulty because it did not allow for the possibility of the degenerate state.[1] The authors mentioned have shown, on the other hand, that bodies with the mass of Jupiter and less have too small an internal pressure to produce degeneracy. Apart from any external heating to which they are subjected, their ultimate fate must be to become ordinary solid bodies at zero temperature.

It remains only to add that, if the proton-proton reaction should be found to be a very much slower reaction than it is believed to be by most physicists, then the alternative possibility that a white dwarf could be a star deficient in carbon and nitrogen, rather than in hydrogen, would have to be reconsidered. There is no evidence in favour of this alternative; strictly speaking, however, it cannot be claimed that the point has been finally settled.

Giant Stars

The feature which distinguishes the giant and super-giant stars from those of the main sequence is the smallness of their mean densities. It has not yet been satisfactorily explained why this minority of all the stars of certain masses and luminosities should differ from the majority in such a fashion.

[1] It was the late Sir Ralph Fowler who first pointed out how the existence of degeneracy provides a solution of the white dwarf problem—a solution which was readily accepted by Eddington.

The empirical mass-luminosity relation is a single relation to which the main sequence stars and the giants appear to conform equally well. Remembering how insensitive is the luminosity of a star to certain features of its energy-generating process, not too much must be inferred from this similarity of behaviour. But it is fairly safe to conclude that the giants owe their energy-generation to processes similar to (and possibly identical with) those responsible for the luminosities of main-sequence stars, and also that the region of a giant star in which the main part of its energy-generation takes place must be in a physical state, especially as regards temperature, not greatly different from that of the central region of a main-sequence star.

We can therefore gain a tentative understanding of the problem, if not of its solution, by picturing a giant as consisting of something like a main-sequence star surrounded by a very extensive tenuous gaseous envelope—so to say, a grossly overgrown atmosphere. This is also suggested by the fact that the trend of the "giant-line" in the Russell-diagram is such that the effective temperature, and consequently the radius, may take enormously different values for approximately the same value of the luminosity. A plausible inference is that the outer part of a giant is more in the nature of an appendage than a fundamentally essential part of its luminosity-producing structure.

Following up this view, we then see from the Russell-diagram that only stars in a relatively narrow range of luminosity can possess such an envelope. This indicates that some very special conditions are required for the existence of the envelope, and this is borne out by the comparative rarity of giant stars. It is further emphasized by the existence of what is called the "Hertzsprung gap" between the giant-line and the main sequence in the Russell-diagram; the gap makes it appear unlikely that the condition for the production of an envelope is incipient in ordinary main-sequence stars.

Astrophysicists have recently found that the effect of assuming a variation of chemical composition through a star can be to give a larger theoretical radius than that for a star of the same mass but of uniform composition. This emerges

from the mathematics of the problem. (The effect cannot be understood from the calculations in the first part of this chapter since they apply only for uniform composition.) Now the thermonuclear generation of energy in a star is practically confined to a relatively small central region where the temperature is greatest. Its tendency is therefore to produce a deficiency of hydrogen in that region, compared with the rest of the star, i.e. to result in a variation of chemical composition through the star. It is usually considered that this tendency is counteracted by the mixing effects of convection and diffusion; this view is supported by the general success of the theory of main-sequence stars when uniform composition is postulated. But it is also considered possible that the very special conditions referred to above may be an inhibition of the mixing processes in the particular stars concerned.

The theoretical increase of radius predicted in this way is, however, apparently too small to account by itself for the observed sizes of the giant stars. Some further effect, such as that of radiation pressure must probably reinforce it. But the problem as a whole is still unsolved. Thus the constitution of giant stars is one of the outstanding puzzles of astrophysics.[1]

[1] (*Added in proof*) More recently published calculations by Hoyle and Lyttleton, and by Li Hen and M. Schwarzschild, show that the assumption of non-uniform composition can in fact, by itself, account in a satisfactory quantitative manner for the observed radii of at any rate some giant stars.

A *variable* is sufficiently defined for our purposes as any star whose luminosity is observed to vary. Some idea of the proportion of variables amongst the stars is given by the fact that about 3 per cent of naked-eye stars are variables. But almost all known variables are stars of large absolute magnitude; so the proportion of variables amongst all the stars must be considerably less than this figure.

The variables form a heterogeneous set of stars as regards both their observable characteristics and also the theories which attempt to explain them. The various phenomena are as yet very imperfectly understood; but there is no doubt that they will in due course shed much light upon problems of stellar structure and particularly upon those of stellar evolution.

Classification. Variables are classified as either *periodic* or *non-periodic*. A periodic variable is one which goes through a regular cycle of changes which is exactly, or almost exactly, reproduced over and over again. The duration of the cycle may be anything from a few hours to a few years. The observed periods do, however, tend to fall into a small number of fairly definite groups; stars with periods less than about 50 days are classed as short-period variables.

The non-periodic variables are further classified either as *irregular variables* or as *novæ*. As the name implies, the fluctuations of luminosity and other changes in the irregular variables possess no recognizable periodicities.

Something like half the known variables have been shown to be periodic. Of the other half, a fairly small proportion are known to be irregular, while the rest have not yet been classified. It is, of course, a much simpler matter merely to detect that a star is a variable than to classify its type of variability.

Cepheid Variables

The majority of the short-period variables have a number of general characteristics in common; those so characterized will be included here under the term *Cepheids* (so called after the typical such star known as δ Cephei). These have been more fully studied than other periodic variables and are the only ones it is proposed to discuss here. A further subdivision into two sets of stars with periods grouped around 12 hours and around something like 7 days, respectively, is important for some purposes but need not concern us here.

Cepheid characteristics can be briefly summarized thus:

(a) They are stars of great absolute luminosity; the Cepheids of smallest absolute luminosity are more luminous than ordinary giants and those of greatest luminosity are more luminous than super-giants. (See Fig. 4, p. 124.)

(b) During each period the luminosity of a Cepheid passes through a maximum and a minimum, the luminosity at maximum being on the average about twice that at minimum.

(c) The maximum is usually sharper than the minimum and the brightening to maximum quicker than the subsequent fading to minimum.

(d) The luminosity changes are attended by changes in the spectrum. The spectrum of a Cepheid at any instant can be assigned to a definite class in the spectral sequence, this class being "earliest" when the star is at maximum luminosity and "latest" about the phase of minimum luminosity. As is also borne out by other evidence, this implies that the luminosity fluctuates over a greater range in shorter than in longer wavelengths.

As an example, δ Cephei varies between class F1 at maximum to G3 at minimum.

(e) The spectral changes are accompanied also by changing Doppler shifts of the spectral lines.

These shifts are such as would be produced if the photosphere of the Cepheid is pulsating radially. If they are interpreted in this way, they imply that

maximum luminosity occurs when the *speed* of expansion is greatest (and not when the radius is greatest), and the minimum about when the speed of contraction is greatest. They also imply that the radius of the star pulsates through about 10 per cent, on either side, of its mean value.

Period-luminosity relation. As might be expected, the various properties of the Cepheids are found to be empirically related to each other. Much the most remarkable relation is that between the period of a Cepheid and its luminosity.

For this purpose, the luminosity means the absolute luminosity averaged through the cycle. When this is plotted against the period the points (each representing the luminosity and period of a single Cepheid) in the resulting diagram all lie very close to a single curve. That is, all the Cepheids having the same period have also (very closely) the same luminosity. The relation is such that the longer the period the greater is the corresponding luminosity.

This famous period-luminosity law was discovered in 1912 by Miss H. S. Leavitt and has become more firmly established by the data which have accumulated since then. Its fame has very little to do with the study of variables for their own sake, but is owed to the fact that it has played an indispensable part in the progressive measurement of the astronomical universe.

If an object is recognizable as an individual star, then its *apparent* luminosity can be measured. If the luminosity is observed to vary periodically then the period can also be measured. If the period is less than about 50 days, then the star is almost certainly a Cepheid. Therefore its *absolute* luminosity may be inferred from the period-luminosity curve. Comparison of the apparent and absolute luminosities then gives the *distance* of the star. Owing to the great luminosity of the Cepheids and the simple character of the observations required, this method can be applied to tremendously greater distances than any other which serves for individual stars and yields results of comparable reliability.

The crucial value of this method is that it can be used to

rest can also be translated into absolute distances. Thus it is the phenomenon of Cepheid variability which has made it possible for astronomers to express the greatest distances in the universe in terms of ordinary units of length.

Period-radius relation. It is found that the period of a Cepheid is approximately proportional to its mean radius. The relation to which all Cepheids approximately conform can be written

$$2R/P = 100 \text{ km./sec.}, \qquad (1)$$

where R is the mean radius of the star in kilometres and P its period in seconds.

Period-density relation. Another important empirical relationship shows that the period of a Cepheid is inversely proportional to the square root of the mean density. Miss C. H. Payne was the first to analyse the observational evidence for this relation.

It is to be noted that the three relations which have been mentioned depend successively less directly upon observation. The form of the period-luminosity relation comes directly from observation by using Cepheids which are known from other evidence to be all at approximately the same distance. The radius used in the period-radius relation has to be got by first estimating an effective temperature or its equivalent. Finally, the density to be used in the period-density relation is obtained from this radius together with a mass estimated from the mass-luminosity law.

Pulsation hypothesis. For obvious reasons, an eclipsing *binary*, regarded as a single object, has a periodically variable luminosity, the period being the period of the orbital motion. If, however, it can be established observationally that the object

eclipsing binary, the variability being attributed to partial eclipses by a much less luminous companion. Modern observational data show, however, that this interpretation is untenable. The various characteristics described above are quite inconsistent with any simple eclipsing effect. But the most fatal objection is that the orbit inferred from the observed velocities and period would be smaller than the stellar components themselves, that is, one component would overlap with or enclose its companion. So even if the velocities have to be ascribed to some kind of orbital motion, the light variation could not be ascribed to eclipses in any ordinary sense.

It was H. C. Plummer who first[1] suggested in 1913 that the observed velocity changes of a Cepheid could be attributed to a periodic contraction and expansion of a single star under the combined action of its own gravitation and the elasticity of its material. The observational investigations of H. Shapley, which led to the formulation of the characteristics we have mentioned, afforded strong support for this view. Incidentally, it was the endeavour to explain the physics of such a process which in 1917 started Eddington on his long series of mathematical studies of stellar structure and so inaugurated the modern period of theoretical astrophysics.

It is certainly the case that the theory of the radial pulsations of a gaseous star can account in a quantitative manner for several characteristic features of Cepheid variability. The theory is not contradicted by any major observational feature (except that, as at present treated, it appears to predict the wrong phase-relation between luminosity and velocity changes), but there remain some such features for which it has so far failed to account, including the most notable of all—the period-luminosity relation.

We shall here confine ourselves to a few simple arguments

[1] In 1879, A. Ritter had, however, suggested that a similar hypothesis might account for the variability of certain stars.

time required for a sound wave to make the same journey through the material.

The speed of sound in a gas is proportional to the square root of the temperature. For a family of stars built upon a single model the temperatures at corresponding points are proportional to M/R according to equation (7) on page 137 (and now ignoring differences of chemical composition). Since the distance of travel of the sound wave considered is proportional to R, it then follows that the time of travel is proportional to $R \div \sqrt{(M/R)} = \sqrt{(R^3/M)}$. This last quantity is in fact inversely proportional to the square root of the mean density. Therefore we have obtained a simple theoretical explanation of the empirical relation between period and density.

Moreover, the empirical relation (1) would then show that the mean speed of sound through any Cepheid is about 100 km./sec.

A standard formula shows that the speed of sound in ionized hydrogen at temperature t is $0{\cdot}166\ \sqrt{t}$ km./sec. We do not know how the temperature varies through a Cepheid. But we do know that the boundary temperature is of the order of 10,000 degrees giving a speed of about 17 km./sec. Also we believe the central temperature of any star to be of the order of 20 million degrees giving a speed of about 740 km./sec. If the speed increases uniformly from the boundary to the centre then its average value according to these figures would be almost 200 km./sec. (This is a standard calculation: it must be remembered that the average speed is not in this case the average of the least and greatest values.)

The difference between the 200 km./sec. given by this extremely crude calculation and the "observed" 100 km./sec. is scarcely significant. What is significant is that the two values are of the same order of magnitude. This means that

the region in which the temperature is less than a million degrees. We therefore conclude that the period of a Cepheid is very largely conditioned by the region which we have called the "envelope" in the case of giant stars (page 161). So we cannot expect to have an exact quantitative theory of Cepheid pulsation until we have a satisfactory theory of stars which possess extensive envelopes.

Comments. Enough has been said to show how the pulsation theory can account quantitatively for some of the chief empirical results. We have seen, on the other hand, that no ordinary "binary" hypothesis will fit the facts. Nevertheless, the fundamental weakness of the pulsation hypothesis is that it has not yet been explained why certain stars should pulsate while the vast majority do not. As against this, the rotation of stars and the possession of companions are both common phenomena. Various attempts have been made to explain Cepheid variability by some combinations of these phenomena different from the ordinary binary hypothesis. These attempts envisage something like a binary star revolving inside a common stellar envelope.

It is found that these alternative hypotheses also require periods of about the observed order and yield the same period-density relation as the pulsation hypothesis. In other respects they have been less fully worked out. But they do serve to show that the case for the pulsation hypothesis is less conclusive than it is usually taken to be. They show too that the root-cause of "intrinsic" variability of an object which looks like a single star may be in the last resort not essentially different from the simplest of all causes of apparent variability, i.e. a mere eclipsing effect.

It thus appears that the phenomenon of Cepheid variability still provides a subject for much further research.

increases by a very large factor and then, less suddenly, returns to about its original value.

Since all known novæ have been very inconspicuous stars before the outburst, they have attracted no particular attention until the outburst has occurred. Indeed, without the aid of photography there was no direct evidence of their prior existence and so it was not unnatural that they should be called "new" stars.

Nowadays, when a nova is observed it is usually possible to identify the star in its pre-nova state on some photograph of the region of the sky concerned. It even happens sometimes that an astronomer has photographed the region for some other purpose within a few days or hours of the nova outburst. In that case it is possible to know something about what the star has, or has not, been doing just before its sudden rise to notoriety. From the moment of its discovery as a nova every possible observation is made until it is judged to have finally relapsed into inaction. Even then a few astronomers make it their business to watch for any signs of restlessness. Thus the knowledge of any nova after the time of about maximum brightness is incomparably greater than that of its previous behaviour.

The existence of two types of nova, *supernovæ* and ordinary or *galactic novæ*, is well established. There is some uncertain evidence for the occurrence also of "dwarf" novæ which need not concern us.

Galactic novæ. There are over a hundred recorded occurrences of galactic novæ, but significant photometric and spectroscopic measurements have been made for only about a dozen of these. About another hundred novæ of this type have been observed also in the nearest external galaxies. Despite an almost baffling variety of detail in the cases which have been fully observed, there are certain leading features which they have in common.

These suffice to show that it is basically the same phenomenon with which we have to deal in all cases. This appears to be true even though it has been found possible to distinguish two or three classes of galactic novæ; so it is unnecessary for us to discuss this classification.

The broad facts about a nova outburst and their immediate interpretation are briefly these: The luminosity of the star increases by a factor of the order of 50,000 in the space of a few hours. It then almost immediately starts to decrease, rapidly at first and afterwards more slowly with numerous fluctuations. Three or four months after the initial outburst, when the luminosity has fallen to a few thousandths of its peak value, its decline becomes steadier and even slower until it reaches about its pre-nova value some two years later.

The rise to maximum luminosity is *as if* the photosphere of the star increases in radius to a few hundred times its original value. When a spectrum has been obtained during this phase, the absorption lines do show velocity-displacements corresponding to such an expansion. After maximum, the spectrum shows a succession of complex changes and for a time it is dominated by emission lines. The structure and displacements of the spectral lines indicate that the star is in process of throwing off shells or jets of gaseous material. In some cases this has been confirmed, after sufficient time has elapsed, by the direct observation of a growing nebulosity near the parent star. Also there are some indications that the so-called *planetary nebulæ* are associated with ex-novæ. These are ring-shaped or shell-shaped distributions of luminous nebulous matter, each surrounding a central star of small absolute brightness but high surface temperature. The star is thought to be the former nova which ejected the material now forming the nebula.

The post-nova state of a star which has experienced an outburst is one of approximately the total luminosity of the Sun but of spectral class about O. This places ex-novæ in the Russell diagram roughly half-way between the main sequence and the white dwarfs. So far as the evidence goes, the pre-nova state is not very different from the post-nova.

reserves. For the heat energy stored within the Sun at any instant is estimated to be equivalent to nearly 50 million years' output, i.e. 5,000 years' output is only about one ten-thousandth of the store. A similar figure is expected to apply roughly to any other star. Consequently no surprise need be experienced if a star is in about the same state before and after indulging in the dissipation of a nova outbreak.

Number of novæ. From the number of novæ actually detected within a specified time, say since the beginning of the present century, and from the circumstances of their detection, it is possible to infer roughly the rate of occurrence of novæ in the whole Galaxy. The estimate is 25 novæ per year. Of these, on the average between one and two are detected by astronomers; the remainder escape detection owing chiefly to their remoteness.

The Galaxy is believed to have existed in much the same state as at present for at least 2×10^9 years. If the estimated rate of occurrence of novæ has been maintained, there must have been about 5×10^{10} nova outbursts in that time. But the total number of all stars in the Galaxy is estimated as not more than about five times this number.

This result shows that the nova phenomenon is probably of considerable significance in the evolution of the stellar system. The extreme possibilities are that a certain set of stars must be prone to nova outbursts and must suffer them many times over during their lives, or else that some considerable proportion of all existing stars have already suffered an outburst at least once in their lives.

Stars which are known to be ex-novæ occupy, as has been stated, a special position in the Russell diagram and it is believed that they occupied about the same position as pre-novæ. This suggests that the great majority of stars are *not* ex-novæ

similar outbursts in 30 years, each resembling a nova outburst, though on a somewhat small scale. A few of the well-known novæ have experienced secondary outbursts some decades after the principal outburst (i.e. "principal" so far as actually observed). Finally there are a few stars for which doubt exists as to whether they should be classed as irregular variables or as recurrent novæ.

Thus the observational evidence tends definitely to favour the conclusion that the majority of stars have never featured as novæ.

The question then arises as to whether the minority which have so featured have done so owing to some external stimulus. Without going into particulars, it may be asserted that there is no evidence of anything special in the environment of known novæ which might provide such a stimulus. Astronomers are now generally agreed that the phenomenon depends only upon the star which exhibits it.

Supernovæ. A supernova is distinguished from an ordinary nova by the fact that its maximum luminosity is, on the average, about 1,000 times that of an ordinary nova, i.e. it is about 50 million times the luminosity of the Sun.

Supernovæ occur much more rarely than ordinary novæ. Indeed, we should not yet have known of them as a class had observations been confined to our own Galaxy. It is their being observed in other galaxies (where a supernova may attain at maximum a brightness comparable with that of the entire galaxy in which it occurs) that has led to their recognition as a class by themselves. Having been so recognized, it is now generally agreed that at least two such novæ have occurred in our own Galaxy in historic times; that discovered by Tycho Brahe in 1572 (and which was visible during daylight for some days), and that recorded in China and Japan in 1054. A good case has been made for believing that the Crab Nebula

of the present time resulted from this latter outburst. Supernovæ probably occur at very roughly the average rate of one per galaxy per century.

In their most general qualitative features, supernovæ resemble ordinary novæ, but quantitatively the happenings are on a "super" scale. This is all that need concern us, though it has recently been found that there may exist more than one class of supernovæ.

Theory of Novæ: Stellar Evolution

It is now generally believed that a normal star is losing energy by the conversion of its hydrogen into helium. Therefore it is suffering a progressive change. But we have had occasion to stress the slowness of this process and to point out that for most purposes the star may be treated as being in a steady state.

Although an ordinary nova outburst apparently leaves the star concerned in much the same state as it was in just before the occurrence, the outburst itself denotes a violent departure from a steady state. Therefore we can scarcely escape the inference that in the nova phenomenon we are actually witnessing some phase in the evolution of a star.

If the luminosity of a normal star is produced by thermonuclear processes converting hydrogen into helium, then (as was explained on page 158) the luminosity is expected not to change to any considerable extent until the hydrogen supply is almost entirely exhausted. In the case of the Sun we saw that this process would require about 10^{11} years, or something of the order of 100 times its estimated past life. But compare this with the case of a more luminous star. A component of the eclipsing binary star V Puppis, for example, has mass about 20 times that of the Sun and luminosity about 4,000 times the Sun's luminosity. So it is consuming hydrogen 4,000 times as fast as the Sun; therefore, if both started as consisting of almost pure hydrogen, the time required for V Puppis to exhaust its supply would be about $20 \div 4{,}000 = 1/200$ of that for the Sun. This is less than the past life of the Sun. By appealing to the mass-luminosity relation, we can conclude

quite generally that a normal star of ten or more times the mass of the Sun cannot be as old[1] as the Sun.

There are plenty of these more massive stars shining quite normally at the present time, and we have to suppose that they were formed very much later than the Sun. But it is then natural to suppose that there must now be some amongst them that are on the verge of having exhausted their hydrogen supply.

Let us then consider the behaviour of a star which has reached this state. The depletion of the energy-sources due to the shortage of hydrogen causes the star to begin to contract under its own gravitation. But this contraction must produce an increase of density and temperature in the star. This increase must in turn cause a speeding up of the thermo-nuclear reactions and so must for a time actually produce an increase in the luminosity. Thus, as the star's capital in the form of unconsumed hydrogen approaches exhaustion, the star becomes more and more profligate in its expenditure. Relatively soon, therefore, the interior of the star must become virtually devoid of hydrogen, and thermo-nuclear energy-generation must also cease.

The star must then proceed to contract and there will ensue a phase during which its luminosity is produced almost entirely by the release of gravitational energy. During this phase of the star's evolution the Helmholtz-Kelvin theory thus comes into its own. But we know from the examination of this theory in the case of the Sun (page 92) that significant contraction must take place in a time that is short compared with the past life of the star.

If no new factor comes into operation, the collapse must proceed indefinitely unless and until it is checked by the material of the star becoming degenerate (page 154). But Hoyle has called attention to a factor which almost certainly takes charge

[1] We may note that this conclusion would be invalidated only if the stars can replenish their hydrogen by accretion from interstellar space. While the possible importance of accretion in various connexions is still an open question, most astrophysicists would consider it unlikely appreciably to affect the line of argument we are now following. The general plausibility of the further conclusions to b described tends to support this view.

of the situation at some stage. This is the ***rotation*** of th and is what we must now consider.

A great deal has been discovered from spectros evidence about the rotation of stars. All that need be q here, however, is the remarkable fact, which appears to b established, that stars of early spectral class (O, B, A and F) possess in general very considerable rotations { equatorial speeds of the order of 100 km/sec. The m stars we are considering belong to such classes.

Now, so long as a star remains a single isolated bod rotational *momentum* must remain constant. This means if the star contracts its *speed* of rotation must increase. In if the contraction is uniform, the equatorial speed wi inversely proportional to the radius.

We know also from common experience that if a bc made to rotate sufficiently fast it will break up. The investig of this phenomenon in the case of astronomical bodie formed the subject of many famous theoretical reseai For our present purpose, however, a very simple crit suffices. The increasing equatorial speed of a contracting must tend to reach so great a value that the correspon centrifugal force would exceed the force of gravity at the su of the star. But this would be an impossible state of affairs the material near the star's equator could not then be ret by the star. It is therefore necessary to conclude that rotat instability must result in the star losing some of its ma before this stage has been reached.

A numerical illustration may be useful. The compone the star V Puppis, already mentioned, has radius about 7 radii. It is of class B1, and the typical equatorial spee rotation for that class is 100 km./sec. (It does not m whether the particular star for which the other figure taken happens to have this speed.) A simple calculation s that, if such a star were to contract uniformly to a radius a one-eighth of the Sun's radius the centrifugal force at equator would exceed the star's own gravitation. Rotat instability would therefore be expected to set in before radius has contracted to that value. It should be added such contraction in itself, though drastic, is not impossib

would result in a mean density which is still only between one-tenth and one-thousandth of the mean densities of white dwarfs.

The star must remain nearly spherical until the centrifugal force at the equator becomes comparable with the force of gravity, i.e. until it is approaching the critical radius. It will then start to bulge at the equator and, with increasing speed of rotation will tend to become lens-shaped. When rotational instability actually occurs, the star will throw off material from the rim of the "lens". The discarded material will be endowed with sufficient speed ultimately to escape completely from the star's attraction.

At the first onset of rotational instability sufficient material must be thrown off to relieve the immediate degree of instability. But this does not arrest the contraction of the parent star. So we must suppose that the process goes on being repeated a number of times.

All this is a brief account of Hoyle's theory of the nova phenomenon. In essence, it is quite simple: a star uses up its normal source of energy; it then proceeds to contract and its speed of rotation increases; a time comes when it can no longer hold itself together in one piece and so some of its matter breaks away; when this is happening, the star is what we recognize as a nova; the process must recur and so the same star will become a nova on several occasions.

Granting only that a star is using up its hydrogen and that it is rotating, the whole process is, moreover, one which not only can but, according to simple physical principles, *must* occur. Accepting this, and knowing that the nova phenomena does occur and does match the theoretical predictions in essential respects, the theory becomes highly convincing.

The theory has not yet been worked out in detail and several features require further elucidation. Thus it has not yet explained why material should be thrown off by a nova in sudden bursts, and it can therefore provide no theoretical estimate of the quantity ejected in a single outburst. Indeed, we might rather expect that the onset of rotational instability would cause matter to stream away in a fairly continuous manner. The behaviour of what are known as P Cygni and

Wolf-Rayct stars indicates that ejection does sometimes take this form.

On the other hand, Hoyle is able by a development of his theory to predict the *total* amount of material which must leave the star before it is finally immuné to any further rotational instability. It reaches such a final state partly because the ejected material carries away some of the troublesome angular momentum, and partly because the remaining material ultimately becomes degenerate and consequently less compressible. Hoyle gives reasons for concluding that the final mass of the star should not exceed about the maximum mass of a white dwarf star as predicted on quite different grounds by Chandrasekhar. This mass is about three times the mass of the Sun.

The theory contemplates the final state of a collapsing star to be in all cases a white dwarf. Hoyle points out, however, that it may not be true, conversely, that every white dwarf has been produced in this way. Current estimates of the white dwarf population give too many such stars for it to be likely that they all originated from novæ.

Relative abundances of the elements. A digression on this topic is now required.

When we discussed thermonuclear reactions we were considering a restricted class of nuclear reactions proceeding under a restricted class of physical conditions. The reactions were those involving the interaction of only two nuclei at a time and, moreover, one of these two had to be a proton (since only reactions in which a proton is captured can release an amount of energy that is significant for stellar energy-generation). The physical conditions were that the chemical composition was prescribed and that the temperature and density were those considered to be appropriate to normal stellar interiors. When the conclusions regarding such reactions are incorporated in the theory of stellar constitution it is confirmed that the physical conditions are of the general character which had been assumed, and, of course, those conditions could then be more accurately computed.

One inference was that reactions the reverse of those considered do not take place at any appreciable rate in ordinary

stellar interiors. The thermonuclear reactions are in fact analagous to "slow" chemical reactions proceeding in conditions of thermodynamic equilibrium. But the proportions of the reagents present at any instant are not, of course, characteristic of that equilibrium. If they were, then each reaction and its reverse would proceed at exactly the same rate and there would be no progressive change and no net release of energy.

If, however, the temperature were something over 10^9 degrees (instead of being of the order of 10^7 degrees), then it can be shown that the reactions previously considered and their reverse reactions would proceed at comparable rates. Therefore, if the reactions are considered as proceeding in some sort of enclosure, there would no longer be the possibility of a slow one-way reaction. Instead, the reactions would proceed until the proportions of the reagents become such that a balance is struck between each reaction and its reverse. Thereafter, both these reactions would proceed at exactly the same rates so that the proportions persist at steady values. These values are then characteristic of the complete thermodynamic equilibrium thus attained.

If, moreover, the density of the material as well as its temperature is also very much greater than that previously considered, then all kinds of reactions involving more than two nuclei at a time will come into operation. These, fortunately, need not be considered in detail. We need only accept the fact that they make possible the formation of any nuclear species not already present. Therefore, if a given mass of material is brought to a given temperature and density, provided these are sufficiently great, every type of nucleus will come into existence. The proportions of the various types will depend only upon the temperature and density and *not upon the types of nuclei which happened initially to compose the material.* As in any other problem of thermodynamic equilibrium, these proportions are calculable solely from the properties of the nuclear structures themselves. The proportions achieved do not depend upon the reactions which bring them into being beyond the mere fact that these reactions are possible at the given temperature and density.

Now a great deal of observational work is being done to

determine as well as possible the relative abundances of the chemical elements (i.e. of the nuclear species) in the universe. Though much remains to be done, a fairly reliable first estimate is now available. To cut a very long story short, it is almost certain that *the existing relative abundances are such as would be achieved in the manner just described at a temperature between* 10^9 *and* 10^{10} *degrees and a density of the order of* 10^7 *grammes per cubic centimetre.*

Merely from the knowledge that the elements are transmutable, we must conclude that the present chemical composition of the universe is determined by the treatment which its matter has suffered in the course of the evolution of the universe. The result just stated suggests very forcibly that the significant treatment has been that the matter was actually at some time (not, of course, all necessarily at the same time) compressed to a density about 10^7 g/c.c. and heated to a temperature about 10^9–10^{10} degrees.

Many astrophysicists have speculated as to when and where such conditions could have been found. No-one now believes that they can be found inside normal stars. Some astrophysicists have therefore investigated the possibility of there having existed in the past one or more "super-stars" having internal conditions of the sort required. This is open to the prime objection that it postulates a former distribution of matter for which we have no observational evidence. But it is open also to a direct physical criticism. It fails to explain how the elements, synthesized under the appropriate conditions, have come to survive in the *existing* distribution of matter in the universe. For, in the course of any orderly evolution from the postulated initial distribution to the present distribution, the proportions of the elements would at any stage be characteristic of the current conditions and not of the earlier conditions, i.e. the elements that were synthesized when the required conditions obtained would not survive in the same proportions in the redistributed material.

Supernovæ and the abundance of the elements. This brings us to Hoyle's highly significant suggestions concerning the mechanism of supernovæ.

In the numerical example on page 176, at the stage when rotational instability occurs, the central temperature of the collapsing star can be estimated to be about 10^9 degrees. Had we, however, assumed an initial speed of rotation, say, one-tenth the value employed, then the critical radius would have been one-hundredth the value calculated and the corresponding central temperature about 10^{11} degrees. The central density would be colossal. But we are not at the moment asserting that the star *will* reach that state—merely that rotational instability alone would not hinder its doing so.

The important point is that, before rotational instability occurs, a stage will be reached when the central temperature is in the range 10^9–10^{10} degrees and the central density of the order of 10^7 g./c.c. Now these are just the conditions required for the synthesis of the elements in their requisite proportions. But that is not all. We know that the speed of the nuclear reactions is exceedingly sensitive to temperature and density. During the collapse of the star, until these particular conditions are attained the reactions are quite insignificant. When the conditions are attained they proceed very quickly and something like statistical equilibrium between the nuclear species results almost immediately.

According to Hoyle's analysis, the subsequent course of events is somewhat complex. Only two features need be mentioned here. The first is that the change in composition of the material leads to a phase in which the collapse of the star becomes catastrophic. This brings the star to a stage when rotational instability comes suddenly into full play with the result that material is thrown off in a tremendously more violent manner than in the relatively gradual collapse producing an ordinary nova outburst. Such is Hoyle's explanation of a supernova.

The second feature is that the ejection of matter is so sudden that the matter is in fact composed of a mixture of the elements synthesized in the preceding phase. That is to say, there is no opportunity for the synthesis to be reversed before dispersal takes place. Thus the sequence of events leads not only to the synthesis of the elements in the required proportions but also to the dissemination through space of the products of the

synthesis. If the products are rapidly dispersed in this fashion they are, of course, transferred to conditions in which no further nuclear transmutations can take place. In this way the theory surmounts the difficulty, stated in the preceding section, of accounting for the survival of the nuclei once they have been synthesized.

Hoyle's achievement is to account for the presence in interstellar space of the elements in the observed proportions. Once this is accomplished it is not difficult to account for their presence in stars and elsewhere in the universe. That is, however, another story.

Summary and discussion. To summarize Hoyle's theory of novæ: We think of any star as consisting initially almost entirely of hydrogen. Throughout its normal life it depends for its luminosity on the conversion of hydrogen into helium. We consider in particular a fairly massive rotating star that has practically exhausted its hydrogen. It proceeds to collapse.

If the original rotation is sufficiently fast, rotational instability operates before there can be any further change in chemical composition and causes the ejection of matter either continuously (Wolf–Rayet stars, etc.) or in bursts (ordinary novæ).

If the rotation is less fast, synthesis of heavy elements occurs before rotational instability can operate; this leads to further rapid collapse followed by the vehement ejection of matter (supernovæ). The ejected matter contains heavy elements synthesized shortly before the sudden collapse.

The implication is that all the heavy elements now found in the universe have been synthesized in this way in the supernovæ that have occurred during the history of the stellar system. A star behaving in this way at the present time is considered to have started life with some heavy elements produced before its birth. But the earliest supernovæ may have been stars which started as pure hydrogen and which owed their normal luminosity to the proton-proton reaction and not to any trace of heavy elements facilitating the carbon-nitrogen cycle.

Finally, what remains of the present star is in all cases considered eventually to form a white dwarf.

It is necessary to admit that there is no generally accepted view on these questions. The details have as yet been insufficiently discussed. But Hoyle's ideas carry considerable conviction because they connect so naturally with what are generally accepted views about stellar structure and energy-generation and afford a coherent explanation of several major observational phenomena. There appears, too, to be no rival explanation that is free from fundamental objections.[1]

If the theory as sketched is substantially valid, it means that an important chapter has been written, at any rate in first draft, in the story of stellar evolution. But it is clearly not one of the early chapters. Thanks to the work of a number of authors it may be claimed that some of these also are now in draft form. But it would take us outside the scope of this book even to summarize them. For the rest of the story is not that of the stars as virtually independent, self-supporting bodies. It is the story of their relationship to the whole material system out of which they have been formed—in fact, the story of the evolution of the Galaxy. This involves the theory of galactic structure which is a subject even vaster than that of stellar structure, though it has been less fully investigated.

. . . . . .

This book has been about just one department of astronomy. It is hoped that the foregoing pages show that certain of the fundamental problems of the subject have been solved in principle.

While astrophysicists may claim that this is so, their investigations show them also, however, the immensity of the tasks still to be undertaken. Even where progress has been most encouraging it still demands confirmation and consolidation by the aid of much fuller physical and astronomical data. But there are many fundamental problems well within the scope of the subject in which little indisputable progress has yet been made; we may instance those of the structure of

[1] So far as supernovæ are concerned, Hoyle's theory has some essential features in common with one advanced by G. Gamow and M. Schoenberg.

giant and super-giant stars, of the phenomenon of Cepheid variability, of the origin of the magnetic fields apparently possessed by certain stars, and of the many "peculiar" stars which depart in unexplained ways from the behaviour of the majority. There are other unsolved problems about which it is not known even if they come within this department of astronomy or belong to a more extended one. Here we may instance those of the origins of double stars, of cosmic rays, and of cosmic "radio noise". The last is an example of a problem whose existence has been revealed only in recent years and we must expect more such entirely new problems in the future. Besides all these fundamental problems, there are innumerable more detailed ones concerning such phenomena as, for example, solar activity.

When astrophysicists came to understand in essentials the constitution of main-sequence stars, they had reached a certain hill-top. But the existence of the lesser unsolved problems which have just been mentioned shows that much of the territory traversed on the way has not yet been adequately explored. The existence of the other unsolved problems of stellar constitution shows that there are several comparable hills yet to be climbed. When, however, astrophysicists contemplate the great problems of the evolution of the stars, of interstellar matter, of systems of stars within the Galaxy, and of the Galaxy itself, they see many mountains. Perhaps they can even now see possible routes up those mountains and a few of their number have indeed ventured some distance along them; but the peaks remain to be conquered. And when, beyond these problems, the still greater ones concerning the constitution and evolution of the whole universe of galaxies are considered, it is realized that even these peaks are only outliers of the real mountain range. The hill upon which we now stand is a very lowly eminence.

BIBLIOGRAPHY

The following books supply some of the general astronomical background:

PETER DOIG, *Outline of Stellar Astronomy* (Hutchinson, 2nd ed., 1947).
D. S. EVANS, *Frontiers of Astronomy* (Sigma, 1946).
F. J. HARGREAVES, *Size of the Universe* (Penguin Books, 1948).
H. SPENCER JONES, *General Astronomy* (Edward Arnold & Co., 2nd ed., 1934).
H. N. RUSSELL, R. S. DUGAN, and J. Q. STEWART, *Astronomy*, Vol. II (Ginn & Co., 1938).

The following books supply general accounts of solar physics:

G. ABETTI, *The Sun* (Crosby Lockwood & Son, Ltd., 1938).
F. W. DYSON and R.V.D. R. WOOLLEY, *Eclipses of the Sun and Moon* (O.U.P., 1937).
G. GAMOW, *Birth and Death of the Sun* (Macmillan, 1941).
D. H. MENZEL, *Our Sun* (Philadelphia & Toronto: The Blakiston Company, 1949).

The following books provide general surveys of problems of astrophysics:

A. S. EDDINGTON, *Stars and Atoms* (O.U.P., 1927).
L. GOLDBERG and L. H. ALLER, *Atoms, Stars and Nebulæ* (J. & A. Churchill, 1947).
F. HOYLE, *The Nature of the Universe* (Blackwell, 1950).

The following books are standard monographs on theoretical astrophysics:

S. CHANDRASEKHAR, *Introduction to the Study of Stellar Structure* (Chicago University Press, 1939).
A. S. EDDINGTON, *Internal Constitution of the Stars* (C.U.P., 1926).
J. H. JEANS, *Astronomy and Cosmogony* (C.U.P., 1928).
S. ROSSELAND, *Theoretical Astrophysics* (O.U.P., 1936).
S. ROSSELAND, *Pulsation Theory of Variable Stars* (O.U.P., 1949).

INDEX


B4